1954

A correlation analysis of bounded sound fields

Glass, Charles Joseph

Massachusetts Institute of Technology

A CORRELATION ANALYSIS OF BOUNDED SOUND FIELDS

Charles Joseph Glass

A CORRELATION ANALYSIS OF BOUNDED SOUND FIELDS

GLASS

t Guard Academy

FILLMENT OF THE

DEGREE OF

ENCE

OF TECHNOLOGY

Signature of Author . . . .
Department of Electrical Engineering, August 23, 1954

Certified by . . . .
Thesis Supervisor

Accepted by . . . .
Chairman, Departmental Committee on Graduate Students

8854
On spine: *

GLASS

1954

THESIS
G457

Letter on front cover:

A CORRELATION ANALYSIS OF BOUNDED
SOUND FIELDS

Charles Joseph Glass

A CORRELATION ANALYSIS OF
BOUNDED SOUND FIELDS

by

CHARLES JOSEPH GLASS

S.B., United States Coast Guard Academy
1950

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
August, 1954

Signature of Author . . . .
Department of Electrical Engineering, August 23, 1954

Certified by . . . .
Thesis Supervisor

Accepted by . . . .
Chairman, Departmental Committee on Graduate Students

A CORRELATION ANALYSIS OF BOUNDED SOUND FIELDS

by

CHARLES JOSEPH GLASS

Submitted to the Department of Electrical Engineering on August 23, 1954 in partial fulfillment of the requirements for the degree of Master of Science.

ABSTRACT

This thesis is an extension of the more comprehensive study made by Kenneth W. Goff (5) of the applicability of correlation techniques to the field of acoustical measurements.

The ability of correlation techniques to separate an acoustical signal at a point into components according to (1) the source (2) the transit time from the source to the point in question and (3) the frequency; was used to measure the transient response of a reverberant room. The results of this correlation analysis were compared with pulse measurement data taken under similar experimental conditions. The results of the two methods were found to agree.

A technique for measuring the directive properties of the wave fronts associated with sound fields was devised and tested. It was shown that the property measured corresponded with the definition of "diffusion." Results of diffusion measurements taken in a reverberant room with various portions of the wall covered with a highly absorbent Fiberglas curtain were presented together with corresponding transient response data.

The above mentioned experimental measurements were paralleled by an explanation of the theory and a sample of the mathematics involved in the correlation analysis of bounded sound fields.

Thesis Supervisor: Richard H. Bolt
Title: Professor of Electrical Engineering

ACKNOWLEDGEMENT

I would like to express my appreciation to Professor Richard H. Bolt who suggested and supervised this thesis. I am also indebted to Doctor Kenneth W. Goff for the instruction and constructive criticism rendered in the course of this research.

TABLE OF CONTENTS

Abstract		i
Acknowledgement		ii
Table of Contents		iii, iv
List of Figures		v, vi
	CHAPTER I INTRODUCTION	
1.1	History	1
1.2	Time Series and Correlation	4
1.3	Time Filtering and Frequency Filtering	9
	CHAPTER II INSTRUMENTATION	
2.1	Introduction	13
2.2	The Analog Correlator	13
a)	The Time Delay	14
b)	The Multiplier	18
c)	The Integrator	21
2.3	Data Gathering Accessories	23
2.4	Accessories for Presenting Results	25
	CHAPTER III TRANSIENT RESPONSE OF ROOMS	
3.1	Introduction	26
3.2	Correlation Analysis of Transient Response	32
3.3	Experimental Measurements of Transient Response	42
3.4	Discussion of Experimental Results	48

CHAPTER IV MEASUREMENTS OF DIFFUSION

4.1	Introduction	53
4.2	Correlation Method of Diffusion Measurement	56
4.3	Some Experimental Results	62
4.4	Discussion of Experimental Results	77
4.5	Suggestions for Additional Work	77

CHAPTER V CONCLUSION

5.1	Summary of Correlation Measurements	79

Appendix
Reprint "Pulse Statistics Analysis of Room Acoustics" by R.H. Bolt, P.E. Doak, and P.J. Westervelt. 82

Bibliography 83

LIST OF FIGURES

Fig. No.		
1.1	Interrelations between time function, cross correlation function, amplitude spectrum, and cross power spectrum.	8
1.2 a)	Fourier transform pair for noise having finite bandwidth.	11
1.2 b)	Fourier transform pair for infinitely wide bandwidth noise.	11
1.2 c)	Fourier transform pair for a pure tone.	11
2.1	Block diagram of the analog correlation computer.	15
2.2	Photograph of the analog correlation computer system.	16
2.3	Block diagram of the magnetic time delay.	17
2.4	Simplified schematic of the squaring circuit.	20
2.5	Photograph of the Ampex magnetic tape recorder with a tape loop.	24
3.1	Block diagram of correlator for analysis in the frequency domain.	33
3.2	Volume enclosed by actual and first order image sources for a rectangular room.	41
3.3	Transient response of Room 10-390-A	44
3.4	Spectrum analysis of loudspeaker voltage and microphone voltage for correlation analysis of transient response.	45
3.5	Plot of transient response of Room 10-390-A for pulse and correlation analysis.	47
3.6	Correlation analysis of transient response of Room 10-390-A - Run 2	51
3.7	Correlation analysis of transient response of Room 10-390-A - Run 3	51
3.8	Correlation analysis of transient response of Room 10-390-A - Run 4	52
3.9	Correlation analysis of transient response of Room 10-390-A - Run 5	52
4.1	Auto-correlation function for loudspeaker, microphone system for various noise voltage spectra.	58
4.2	Photograph of microphone array oriented at 0 degrees in sound field of loudspeaker.	60

Fig. No.		
4.3	Photograph of microphone array oriented at 90 degrees in sound field of loudspeaker.	61
4.4	Plot of values of the auto-correlation curve vs accurately measured values of time delay and calculated values of θ.	63
4.5	Diffusion measurements for anechoic space.	64
4.6	Diffusion measurements of experimental studio.	66
4.7	Transient response of experimental studio.	67
4.8	Diffusion measurements of plaster room with bare walls.	69
4.9	Transient response of plaster room with bare walls.	70
4.10	Diffusion measurements of plaster room with one wall covered.	71
4.11	Transient response of plaster room with one wall covered.	72
4.12	Diffusion measurements of plaster room with two walls covered.	73
4.13	Transient response of plaster room with two walls covered.	74
4.14	Diffusion measurements of plaster room with three walls covered.	75
4.15	Transient response of plaster room with three walls covered.	76

I INTRODUCTION

1.1 HISTORY

The word "correlation" implies the act of seeking relationships between two discreet quantities or functions of time or more properly, two ensembles of events. The concept is elementary. For that reason, it is impossible to pin-point some example and cite it as being the first instance of the application of correlation. It is sufficient to say that the process is one which owes its development to man's innate curiosity coupled with his ability to reason.

Historically, correlation is a very important process. When man became dissatisfied with the mere recording of events and first attempted to explain them in terms of a preceding series of incidents, he was resorting to a form of correlation.

Correlation, like many other processes involving basic concepts, lends itself to concise and explicit expression by mathematical symbols. The translating of the process from words to mathematical symbols more or less parallels the evolution of the process from an art into a science.

The mathematical expression of the correlation process was a natural outgrowth of the development of the field of statistical methods. Before 1600, no

mathematical conceptions of probability were recognized. Guilford, in his book on psychometric methods[1] says, "Gamblers had speculated much concerning games of chance when it came time to consider their losses and gains, particularly their losses. They even attempted to interest mathematicians in their problems, though with small success. ... The seventh century saw the beginning of serious interest in the mathematics of chance. ...

"Bernoulli (1654-1705) published the first book to be entirely devoted to the subject. DeMoivre (1667-1754) may be credited with the discovery of the normal distribution curve at about 1733. From that time on, interest was aroused among astronomers as well as mathematicians. By 1812 Laplace (1749-1827) had written what is considered the greatest single work on probability. In it he gave proof of the method of least squares.

"It was Gauss (1777-1854) who demonstrated the great practical value of the normal curve, showing how it applied to the distribution of measurements and to errors made in scientific observations. It was he who devised the fundamental methods of computation of means, probable errors, and the like. ...

"The application of the normal curve and elementary statistical methods to biological and social data must be

[1]Reference numbers refer to the Bibliography.

attributed first to Quetelet (1796-1879), royal astronomer to the King of Belgium. ... Sir Francis Galton (1822-1911) in working on the problems of human heredity found that the normal curve and its simpler applications were inadequate. He invented a number of additional statistical tools, among them the method of correlation."

In 1888, eleven years after Galton's first publication of the concept developed in terms of the ideas of regression lines, the term "correlation" first appeared in print. Pearson, Edgeworth, and Weldon further developed the mathematics of correlation to a point where it was generally adopted by the statisticians in the fields of economics, sociology, biology, astronomy, meteorology and a few other physical sciences. Held, in his book on statistics[2] reviews the various applications of the theory of correlation. As an addendum to his treatment of the subject, Held compiled a bibliography which is excellent in its scope of the subject of correlation.

Perhaps the most important advance in the field of statistics has been the mathematical exposition of the relationships between the field of communication engineering and statistical treatment of time series. The rigid mathematical proofs of Norbert Wiener[3] not only put the process of correlation on a firm mathematical foundation, but also tended to accelerate further advancements in the field of communication engineering.

1.2 TIME SERIES AND CORRELATION

Wiener defines time series as "Sequences, discrete or continuous, of quantitative data assigned to specific moments in time and studied with respect to the statistics of their distribution in time. They may be simple, in which case they consist of a single numerically given observation at each moment of the discrete or continuous base sequence; or multiple, in which case they consist of a number of separate quantities tabulated according to a time common to all."

The application of statistical methods to communication[4] engineering requires a model of an explicit function of time whose properties can be described in terms of its:

1. mean value
2. mean-square value
3. power spectrum
4. correlation function

The function of time, besides being capable of description in terms of a sufficiently complicated set of probability densities, must also be a stationary process; that is, one in which none of the probability distributions which describe the process change with time.

It is necessary at this point to define the statistical quantities associated with the measurement of random functions of time such as the thermal voltage of vacuum tubes.

The mean value: If $\rho(V)\, dV$ is the probability that $V(t)$ lies between V and $V + dV$ then the mean value of a noise is

$$\langle V(t) \rangle = \int_{-\infty}^{\infty} V \rho(V)\, dV$$

under the restriction that

$$\int_{-\infty}^{\infty} \rho(V)\, dV = 1$$

The mean-square value is

$$\langle V^2 (t) \rangle = \int_{-\infty}^{\infty} V^2 \rho(V)\, dV$$

where the notation $\langle \ \rangle$ means the statistical average, as computed from the probability densities which describe the ensemble.

The time-averaged mean value of a noise voltage $V(t)$ is

$$\overline{V(t)} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} V(t)\, dt$$

The mean-square time-averaged value is

$$\overline{V^2(t)} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} V^2 (t)\, dt$$

A random function may be assumed to vanish outside of the time interval $-T/2 \leq t \leq T/2$ in which case

$$V(t) = \int_{-\infty}^{\infty} A(f)\, e^{j2\pi f t}\, df$$

where $A(f)$ is a complex function and equals the voltage spectrum of $V(t)$. With the aid of Parseval's theorem:

$$\int_{-\infty}^{\infty} V^2(t)dt = \int_{-T/2}^{T/2} V^2(t)dt = \int_{-\infty}^{\infty} |A(f)|^2 \, df$$

we may obtain an expression for the total (finite) power of a random function:

$$\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} V^2(t)dt = \overline{V^2(t)} = \int_0^{\infty} W(f) \, df$$

where $W(f)$, the power spectrum, is defined as

$$\lim_{T\to\infty} \frac{2\,|A(f)|^2}{T}$$

The auto-correlation function $\Phi_{11}(\tau)$ of a random function is

$$\Phi_{11}(\tau) = \overline{V(t)\;V(t-\tau)} = \langle V(t)\;V\,(t-\tau)\rangle$$

Similarly, the cross-correlation function of two random functions is

$$\Phi_{21}(\tau) = \overline{V_1\,(t)\;V_2\,(t-\tau)} = \langle V_1\,(t)\;V_2\,(t-\tau)\rangle$$

The link between the correlation function and the other statistical quantities used to describe random function has its basis in Wiener's theorem: <u>the auto correlation function is the cosine Fourier transform of the intensity spectrum</u>.

The general relationships between the function of

time, the spectral density, the power spectrum, and the correlation function are shown in a schematic devised by Kenneth W. Goff[5], Fig. 1.1.

The results obtained by the correlation analysis of acoustical problems are similar in some respects to results obtained by both steady state methods and by pulse methods. From a rudimentary knowledge of the nature of speech and musical sounds, one gets the intuitive feeling that, since the sounds normally encountered in acoustical problems resemble pulsed wave trains, stressing the pulse-like aspects of correlation methods would yield significant results. However, such a supposition is not sufficient to justify the neglect of the steady state aspects of the correlation process. The ability to apply a degree of frequency filtering in the course of computing the correlation function has several advantages. Paramount among these several advantages is the fact that much of the present body of knowledge on acoustical problems is based on tests made with steady state, frequency filtered noises. This fact and the consequent advantages gained by the ability to compare results with those obtained by more classical methods make it desirable to compromise the time filtering properties of correlation methods.

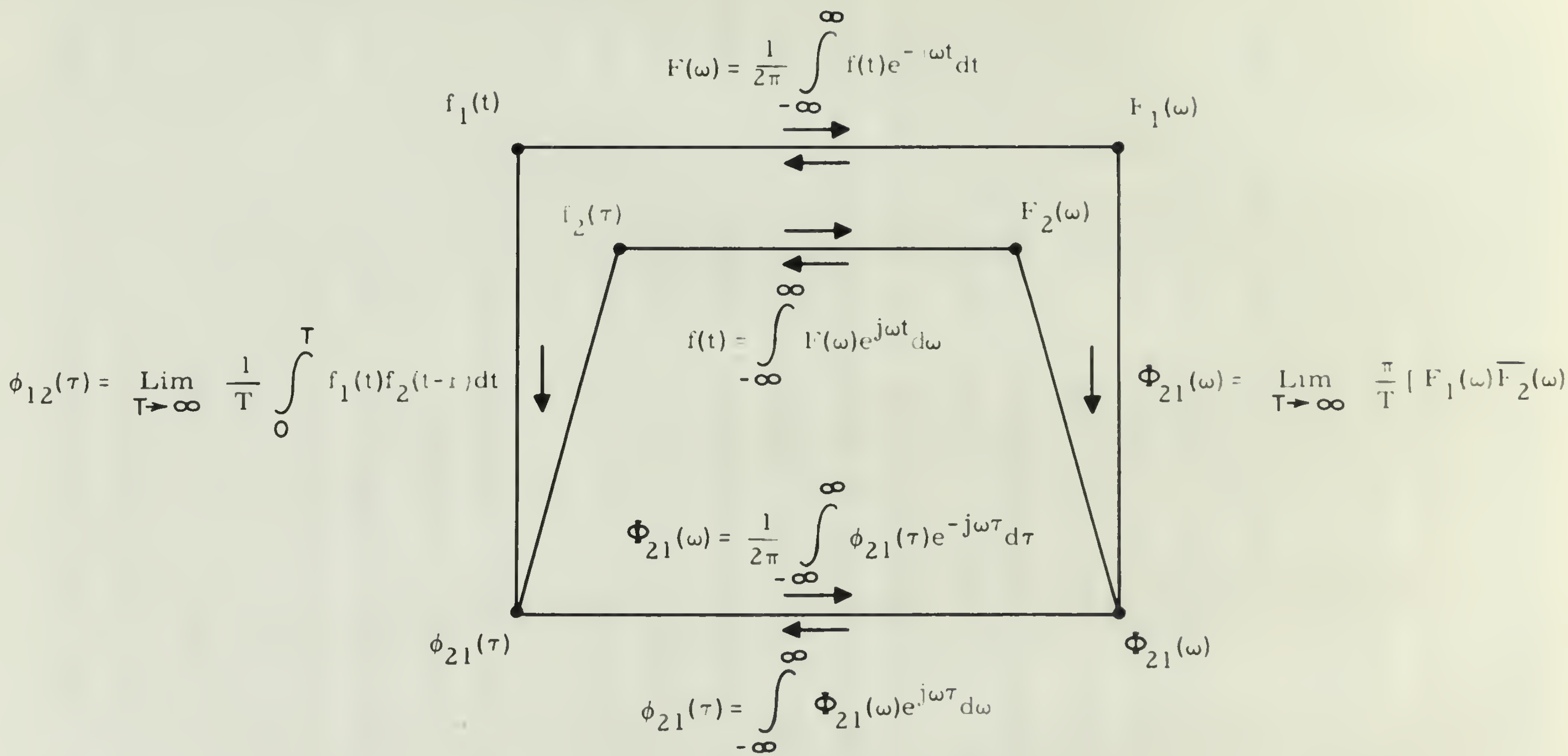


Fig. 1.1 Interrelation between two time functions and their amplitude spectra, cross power spectrum and crosscorrelation function.

1.3 TIME FILTERING VS FREQUENCY FILTERING

Many of the applications of correlation techniques in the field of acoustics depend upon the ability of the method to separate signals into components according to their transit time from a source.

For instance, suppose two microphones were placed in the sound field of a point source radiating random noise into free space. If the microphone outputs were mixed, and this combined voltage were cross-correlated with the loudspeaker input voltage, the following effects would be noticed: At a delay time corresponding to the transit time from the loudspeaker to the microphone closest to the source, the correlation function would exhibit a peak value. At a delay time corresponding to the transit time from the loudspeaker to the second microphone, another peak would appear in the plot of correlation function vs time delay.

If the microphones were placed too close together, the resultant plot of correlation function vs time delay would be distorted. The smallest difference in transit time which produces an undistorted main peak in the correlation function corresponds to the resolving power of the system. This minimum transit time may be of the order of one period of the geometric mean frequency of the bandwidth of the random noise used; and the corresponding

resolving power of the system is equal to the product of this minimum transit time and the speed of sound in the air.

Since the auto-correlation function and the power spectrum of a stationary random time function are Fourier transform pairs, they exhibit certain characteristic inverse spreading properties. For example, a frequency spectrum, $F(\omega)$, flat from 0 to $\pm \pi/\tau$ has as its transform, a function of time, $f(t)$, having the general shape of $\frac{1}{\tau}\frac{\sin \frac{\pi}{\tau} t}{\frac{\pi}{\tau} t}$ which has one predominant peak and axis crossings at $\pm \tau$, $\pm 2\tau$, $\pm 4\tau$, ... , Fig. 1.2 (a). As the bandwidth of $F(\omega)$ increases, the axis crossings of $f(t)$ occur at progressively smaller values of time. The limiting case is an $F(\omega)$ equal to 1. The Fourier transform of such a spectrum is the so called "impulse" or "delta function" which is a function of time having an infinitely small duration, an infinite amplitude and an area of 1 under the curve; Fig. 1.2(b). Conversely, as the bandwidth decreases, the peaks in the function $\frac{\sin x}{x}$ tend to equalize. In the limiting case, $F(\omega)$ is composed of one frequency $(\pm \omega_0)$, and the inverse Fourier transform $f(t)$ is cosine $(\omega_0)\, t$; Fig. 1.2(c).

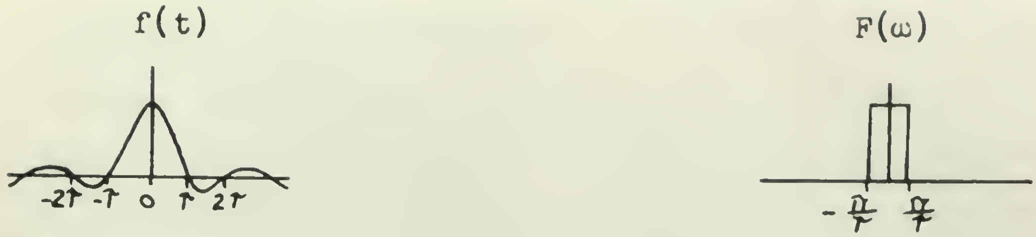


Figure 1.2(a) Fourier transform pair for finite bandwidth noise.

Figure 1.2(b) Fourier transform pair for infinitely wide bandwidth noise.

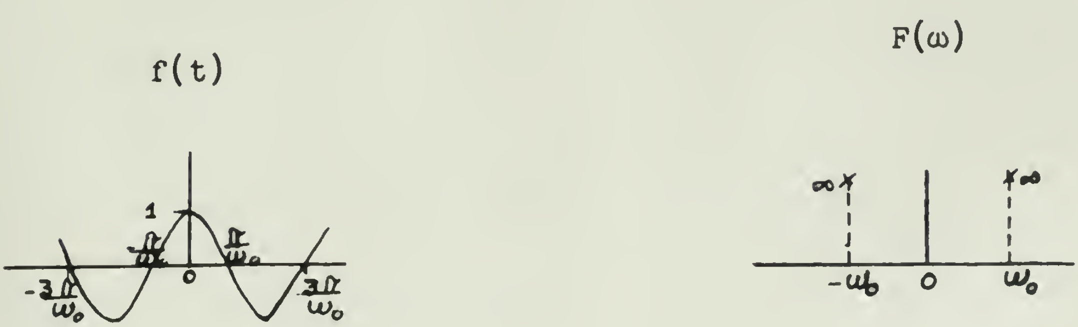


Figure 1.2(c) Fourier transform pair for a pure tone.

From the above discussion and illustration it is obvious that the prerequisites for high resolution in time are diametrically opposed to the requirements for high frequency resolution. It is equally clear that most acoustical measurements made by correlation methods require a compromise between the frequency and time filtering properties of the process.

For instance, in measuring transmission loss of panels, one generally desires to know how this property varies with frequency. Thus the bandwidth of the test signal must be cut to a degree commensurate with the frequency resolution desired. This increase in frequency resolution must not be so great as to create confusion between the successive peaks caused by the signal transmitted directly through the panel and those peaks caused by signals arriving at the test point by different paths.

II INSTRUMENTATION

2.1 INTRODUCTION

The subject of instrumentation must include consideration of the methods of (a) gathering data, (b) reducing data, (c) presenting and recording results.

Equipment for gathering data may include such items as electronic noise voltage generators, loudspeakers, microphones, and in a case where the reducing equipment is not portable, recording instruments. A reducing system is generally composed of a number of separate electronic units each of which performs a specific process. Division of the reducing process into a set of sub-processes lends versatility to the system and also facilitates testing of the components. The end product of the reduction process may be presented in various ways. If the reduction system is electronic, the results may be presented by means of meter readings or oscilloscope traces.

If a reducing system is already in existence, then the equipment for gathering data and presenting results must be selected in such a way that the best use will be made of the features built into the reducing system.

2.2 THE ANALOG CORRELATOR

As implied by the equation defining the cross-correlation function

$$\varphi_{21}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_0^T f_1(t)\, f_2(t-\tau)\, dt;$$

a device capable of computing this function must perform the operations of (a) delay, (b) multiplication, and (c) integration. If the functions of time were electrical voltages such as those produced by a microphone in a sound field, the application of these voltages directly to the input terminals of an electrical analog computer designed to perform the required operation would yield the correlation function directly as an output voltage which could be measured and recorded by properly selected instruments.

The block diagram of such an analog computer designed and developed by Kenneth W. Goff[5] is shown in Fig. 2.1. A photograph of the correlator is shown in Fig. 2.2. For the purposes of analysis, the computer is composed of three separate units

1. time delay system
2. multiplier
3. integrator

Each of these units may be operated independently of the remaining components.

2.2 (a) THE TIME DELAY

The time delay mechanism employed in the computer being considered consists of an electro-mechanical device, Fig. 2.3. As shown in the block diagram, the two signal voltages being used are fed into separate channels each

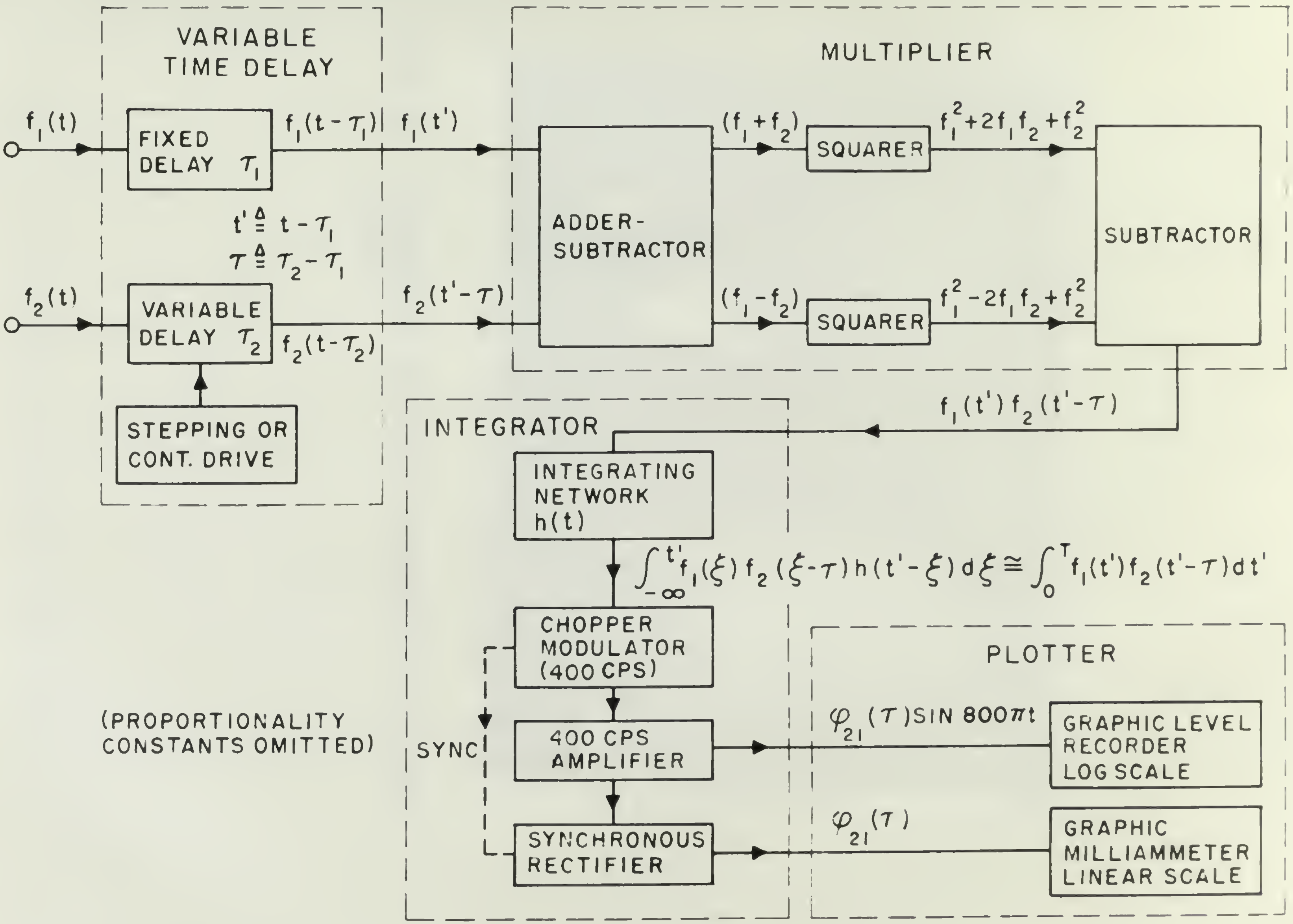
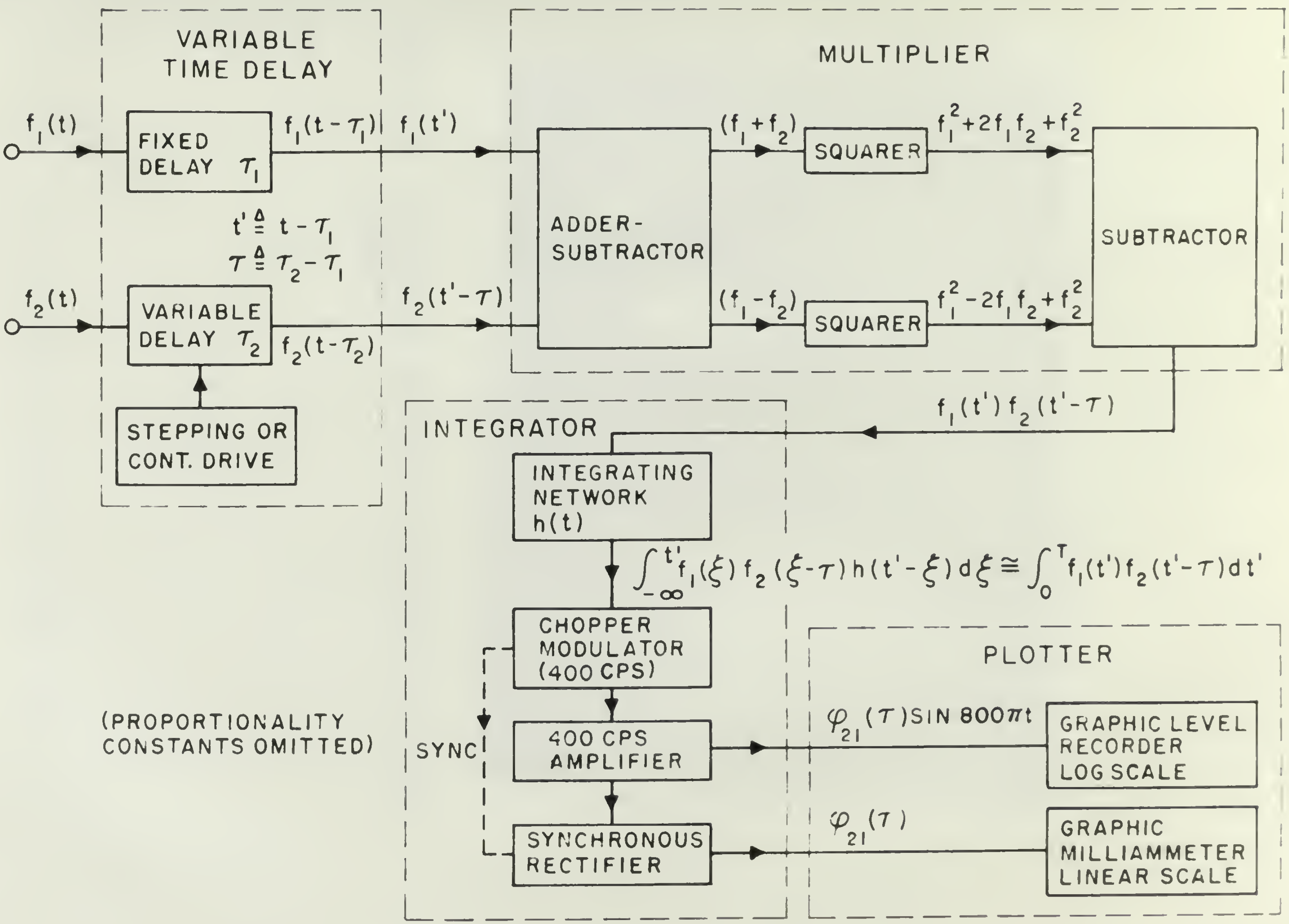


Fig. 2.1 Block diagram of the correlator.



$$\int_{-\infty}^{t'} f_1(\xi) f_2(\xi-\tau) h(t'-\xi)\, d\xi \cong \int_0^T f_1(t') f_2(t'-\tau)\, dt'$$

Fig. 2.1 Block diagram of the correlator.

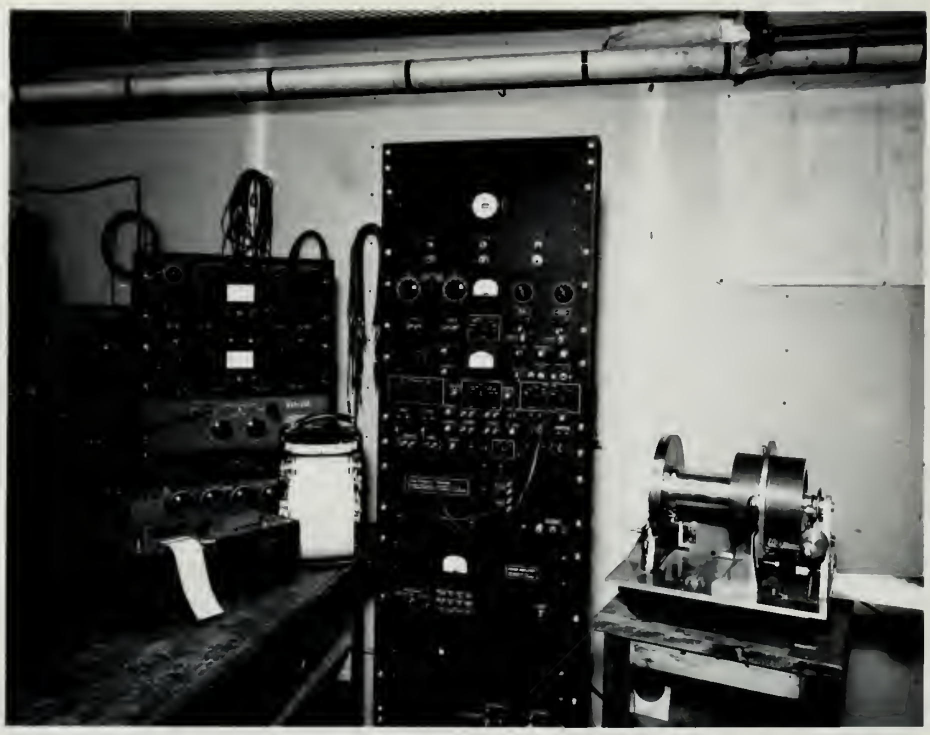


Fig. 2.2 Photograph of analog correlation computer system.

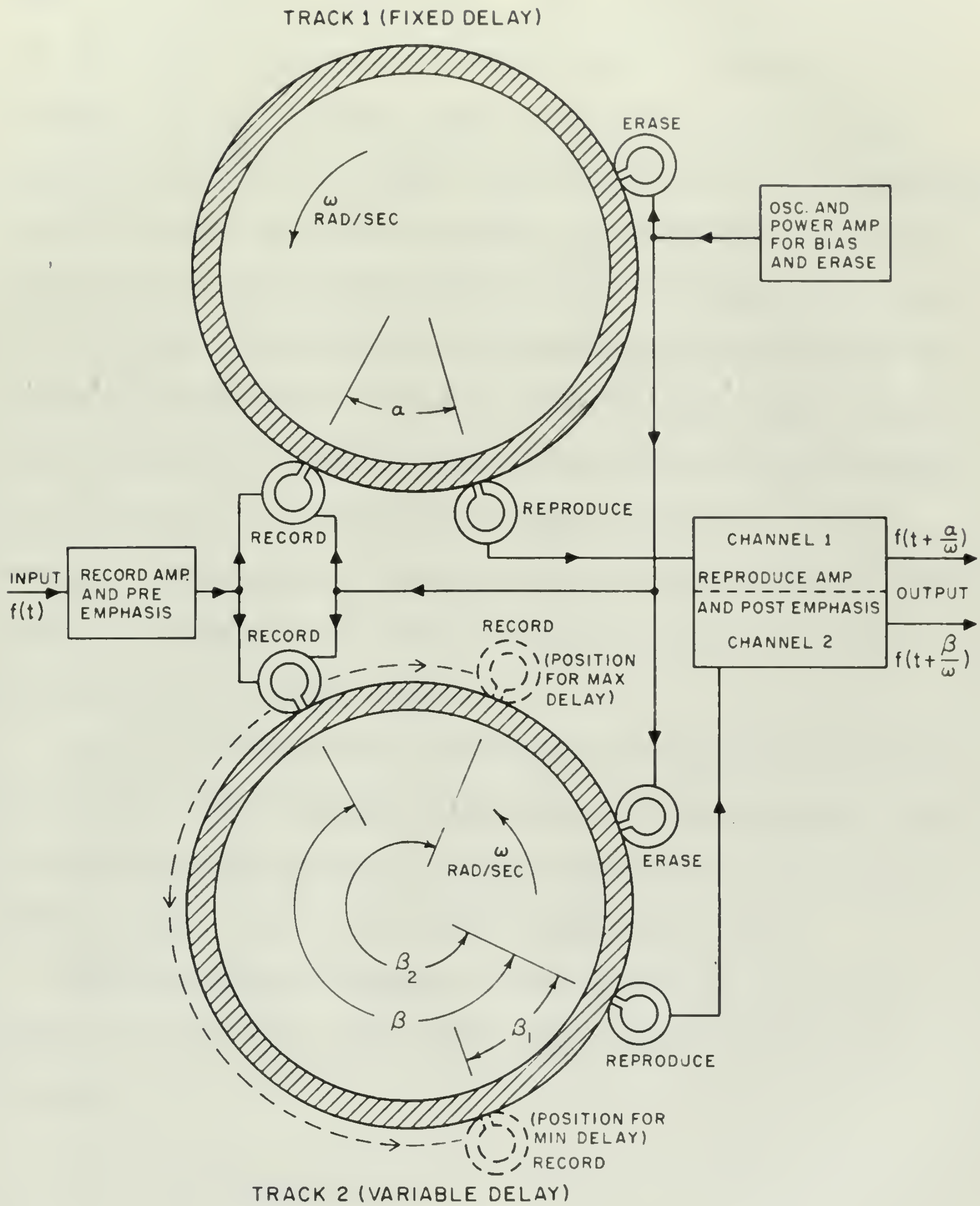


Fig. 2.3 Block Diagram of Magnetic Time Delay

having its own variable gain amplifier, magnetic recording head, magnetic recording track, playback head, erase head, and playback amplifier. The record head of channel one is fixed, whereas the record head of channel two is movable. If the transit times from the record to the playback heads for a point on the surface of the recording drum are equal for both channels, then no relative time delay will be introduced between the two input voltages. If the transit time between record and playback heads is greater on channel two than on channel one, then a relative time delay equal to the difference in transit times is introduced. For identical input voltages to each channel, the output of channel one would be f(t) and the output of channel two would be $f(t-\tau)$ where τ is the difference in transit time.

This relative time delay between channels one and two can be varied continuously from -15 milliseconds to 190 milliseconds by means of gear train and chain drive system powered by a low speed synchronous motor.

The frequency response of the system is flat to within $\pm$ 2 db over a frequency range from 100 cps to 10 kcps.

2.2 (b) THE MULTIPLIER

The multiplier component of the correlator is a quarter-difference-squaring device which reduces the

process of multiplication to that represented by the relation

$$f_1 f_2 = \frac{1}{4} \quad (f_1 + f_2)^2 \ - \ (f_1 - f_2)^2$$

The process of producing a voltage proportional to the square of $(f_1 + f_2)$ and $(f_1 - f_2)$ is performed by a squaring circuit consisting of two 6B8 pentodes operated with grids in push-pull. For operation in the region where the transconductance and grid voltage are lineally related, the plate current for one tube can be approximated by

$$i_b = a + be_g + ce_g^2 + de_g^3$$

For the two tubes operating in parallel with a common plate load as shown in Fig. 2.4, the current through the common load is the sum of the plate currents

$$i_L = i_{p1} + i_{p2} = 2a + 2ce_1^2$$

The output voltage

$$e_o = E_{bb} - i_L R_L = a + Be_1^2$$

If two squaring circuits are operated in push-pull and e_1 is made equal to $(f_1 + f_2)$ in one squaring and equal to $(f_1 - f_2)$ in the other squaring circuit,

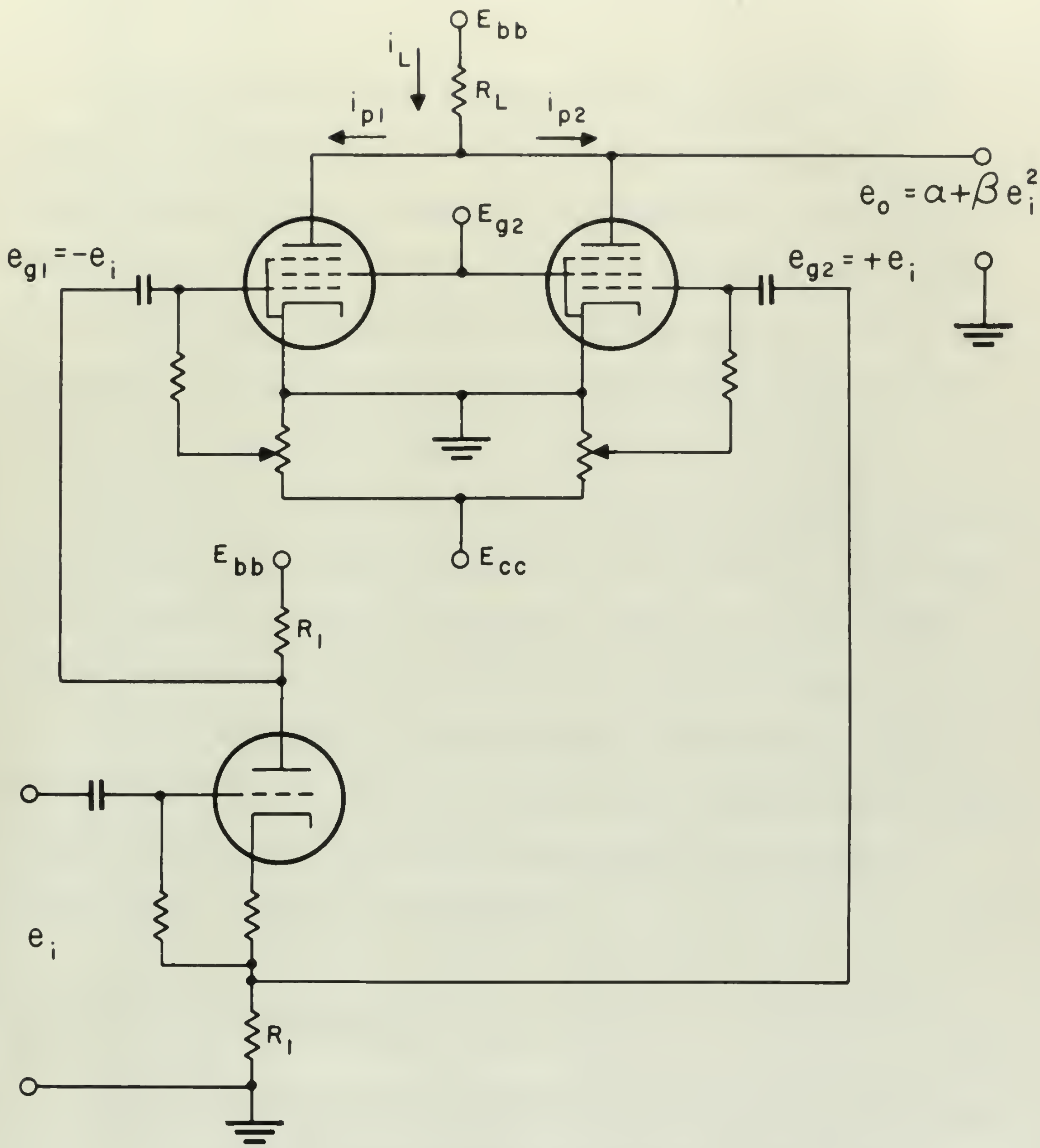


Fig. 2.4 Simplified schematic of the squaring circuit.

then the output voltage will be proportional to the product $f_1 f_2$.

Variability in the performance of the correlator generally results from unbalanced conditions in the multiplier circuits. When the multiplier is operating correctly, the output voltage should be zero when there is no input and when there is only one input. These balanced conditions correspond to the requirements that $0 \times 0 = 0$ and $0 \times 1 = 0 = 1 \times 0$.

The balance for $0 \times 0 = 0$ is made by adjusting the dc bias of the four 6B8 tubes for equal quiescent plate currents whereas the balance for $0 \times 1 = 1 \times 0 = 0$ is made by adjusting the ac levels applied to each 6B8 tube.

Much attention was focused upon the design requirements necessary to reduce dc drift to a degree that would permit continuous operation for periods up to two to three hours without adjustment[5].

2.2 (c) THE INTEGRATOR

Integration in the computer is accomplished by means of either a low pass filtering system or a stepping integrator. For acoustical measurements, a simple RC low pass filtering circuit is satisfactory. The correlator has several such circuits with RC times of 0.5, 1.0, 2.0, 4, 8, and 16 seconds. Selection of the proper RC time is another important factor governing optimum use

of the equipment.

The dc output of the RC circuit is amplified and then placed across the input of a dc-ac chopper which converts the dc voltage into a proportional 400 cps signal. This 400 cps signal is then amplified again and placed across the output terminals of the integrator.

The 400 cps signal may be measured directly with a vacuum tube voltmeter or an oscilloscope, or it may be rectified and recorded by means of a dc graphic ammeter. A linear plot such as is obtained on a graphic ammeter is extremely useful since it represents the shape of the computed correlation function. Information regarding the relative amplitude in decibels of successive peaks of the correlation curve is desired. Such information may be obtained by means of a logarithmic graphic level recorder which records the correlation curves directly in db. When used together, the linear and logarithmic plots give a continuous, detailed, visual record of the shape and magnitude of the correlation curves as a function of time delay.

All of the information presented in the foregoing discussion of the analog correlation computer has been taken from the doctoral thesis of Kenneth W. Goff[5]. Only those points which are considered necessary for a basic understanding of the computing process are presented here. For a more detailed treatment of the

design requirement and the performance characteristics of the computer, the reader is referred to either reference 5 or to the recently completed, unpublished manual "Instruction Manual for Analog Correlator" by Kenneth W. Goff, August 1954, M.I.T. Acoustics Laboratory.

2.3 DATA GATHERING ACCESSORIES

The accessories selected for gathering data are as follows:

1. The noise source consisting of an electronic noise voltage generator, a power amplifier, and an Altec 728B loudspeaker in a closed box baffle.
2. The microphones used were Altec Lansing type 21 BR-150 with 157-A bases and P518A Altec Lansing power supplies. The response curves for these microphones are flat to within ± 2 db for a frequency range of from 100 cps to 10 kcps. The voltage amplifiers used with these microphones were wide band amplifiers.
3. Since the reducing equipment was not portable, provision for recording data was required. An Ampex model 350-2 twin track magnetic tape recorder (Fig. 2.5) was used for this purpose. The particular apparatus used had a frequency response flat to within ± 2 db over a range of

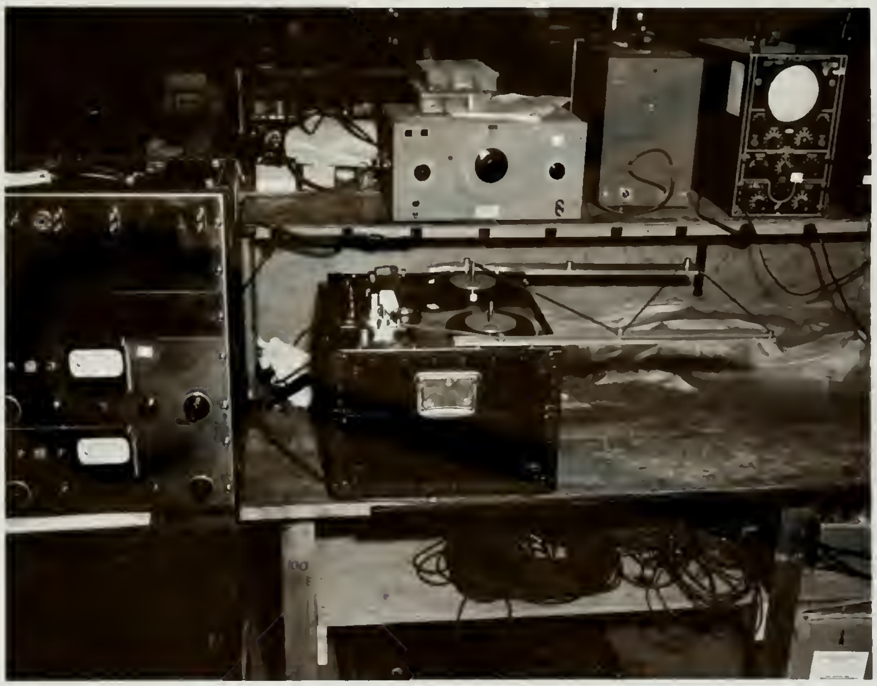


Fig. 2.5 Photograph of Ampex magnetic tape recorder with a tape loop.

150 cps to 15 kcps. This equipment was portable and yet capable of reproducing a relatively distortionless replica of the input voltage.

2.4 ACCESSORIES FOR PRESENTING RESULTS

The correlation function as computed by the analog correlator is presented in the form of a 400 cps voltage and a rectified dc voltage. It is therefore possible to measure and record the correlation function by means of a variety of electrical measuring instruments.

The accessories used for measuring voltages are:

1. Model 300 Ballantine vacuum tube voltmeter.
2. Type 208 DuMont Cathode Ray Oscillcscope.

Permanent records of the output voltage were made by means of:

1. Model AW linear scale Esterline Angus Graphic Milliammeter.
2. Brüel and Kjaer log scale Graphic Level Recorder.

III. TRANSIENT RESPONSE OF ROOMS

3.1 INTRODUCTION

Many studies have been made on the subject of sound fields in rooms. Generally these studies have been based on one of two distinct approaches to the problem of room acoustics:

a) a study of the effect of rooms' boundaries on the sound field associated with steady state pure tone signals.

b) a study of the effects of rooms' boundaries on the sound field associated with transient signals.

Therefore, in attempting to obtain a complete quantitative assessment of the effects of boundaries on a sound field, one might:

a) excite the room with a non-directional steady state sinusoidal signal and map the amplitude and phase relations of the sound pressure for

   1. all frequencies
   2. all possible observer positions in the room
   3. all possible source positions in the room

b) excite the room with a pulse of extremely short duration and measure the relative amplitude and direction of propogation of the associated wave fronts for

1. all time
2. all observer positions
3. all source positions.

If one has all of the data associated with either one of the above mentioned experimental approaches, he can mathematically deduce all of the information for the other method. For example, one can synthesize the pulse used in terms of a Fourier series of sinusoidals. Then for each frequency component one can pick the relative amplitude and phase angle of the sound pressure at any point in the room. By weighting these amplitude and phase measurements in accordance with the Fourier analysis of the pulse and then adding the effects for each frequency, the complete response of the sound field to a transient pulse can be deduced.

On the other hand one can visualize a model of a room excited by a pulse. In such a model, the walls of the room would be eliminated and their effect on the sound field would be simulated by an infinite multidimensional array of image sources, each of which is located in a cell of the same size and shape as the actual room. Each image source is positioned so that the transit time of the signal from the image source to the actual observer's position equals the transit time for the signal arriving at the observer's position by a particular sequence of reflections from the room's boundaries.

With this model, one can now specify that each nondirectional source is in phase and is radiating a free traveling sinusoidal signal. By adding up the contribution from each image source, one can map the amplitude and phase relations of a room excited by a steady state source.

The above discussion assumes that the boundaries of the room are infinitely hard so that no phase shift or decrease in amplitude occurs when a signal is reflected from the walls. This assumption simplifies the problem, but the resulting calculations do not conform with reality. It is necessary therefore to consider these phase shifts and attenuations occurring at the boundaries of the sound field. In the transient treatment of sound fields one may account for the boundary effects by assigning to each bounding surface an impulse response $H(\omega)$. In studying the steady state behavior of bounded sound fields an absorption coefficient and a phase shift are assigned to each wall.

Any such complete analysis of a bounded sound field would be extremely tedious. A desire to be practical would incline one to restrict the analysis to a degree commensurate with the amount of detail necessary for the proper utilization of the room. In most rooms we are interested in producing an acoustical environment conducive to good audio communications. Therefore we may generally

restrict our consideration to the audible frequency range. Audio communications will generally fall into one of three categories.

1. two way vocal communications between individuals
2. vocal communications between one speaker and an audience
3. instrumental communications between musicians and an audience

Traditionally, each type of communication mentioned above has associated with it a certain physical arrangement of source and observer. If we accept these traditional arrangements as being characteristic of the type of communication to be used in the room then we may further limit our analysis of a bounded sound field to those source and observer positions of practical interest.

For practical purposes we may further restrict our analysis by considering only those types of sounds characteristic of musical or vocal sources. Such simplifications further restrict both the frequency composition of the testing signals and pulse duration and pulse intervals for transient signal testing.

Up to this point we have been discussing the two general approaches to the study of room acoustics and the various simplifications that may be employed in this study. These two approaches have been quite thoroughly explored and the resulting body of data coupled with

subjective evaluations of the acoustical environment in rooms has led to the conclusion that variations in both the steady state characteristics and the transient characteristics of sound fields account for differences in the acoustical quality of a room.

W.C. Sabine[6] recognized the importance of the transient behavior of sound fields and attempted to relate it to a subjective evaluation of the quality of rooms. As a measure of transient response, Sabine defined a "reverberation time" <u>the length of time for the mean square pressure of a suitably chosen distribution of sound waves to diminish to one millionth of its original intensity</u>. Reverberation time as defined by Sabine is a measure of the rate of decay of the sound pressures caused by a steady state sinusoidal signal which is suddenly discontinued. Various criteria based on reverberation time vs the size of a room which is to be used for a specific application, and reverberation time vs frequency have been devised[7]. At present, design curves of this type offer the most satisfactory approach to the construction of rooms having good acoustics and the treatment of rooms having poor acoustical characteristics. However, it is recognized that acoustical design criteria other than reverberation time must be considered in the correct design of rooms. For instance, one research group[8] suggested that the first

20 db of decay of a sound in a room is of primary importance in differentiating between two rooms which have approximately equal over-all reverberation times.

C.A. Mason and J. Moir[9], in the course of their studies of the acoustics of theaters by means of pulses or "tone bursts," intimated that the time and amplitude distribution of reflected tone bursts might serve as an aid in evaluating the acoustical quality of rooms.

In the introduction to "Pulse Statistics Analysis of Room Acoustics" the authors state, "... These facts indicate that ... the first few reflections are primarily responsible for certain important features of the acoustical character of rooms ... It should therefore be worthwhile to study this 'short term' transient response by a method of images in which all of the wave properties of the image sources can be considered (i.e. where the assumption of an incoherent source is not made). Further, if an image array satisfying the boundary conditions can be found, one should be able to treat this array statistically and thus obtain the long term average transient response as well ..."

As indicated by the above quotation, some method of measuring the transient response of rooms is required for an evaluation of the effects of boundaries on a sound field. However, some estimate of the steady state behavior of rooms is equally desirable. One experimental method of measuring the transient response of rooms has been

suggested which seems to combine some of the properties of both steady state methods and pulse methods[5]. This method employs correlation techniques.

3.2 THE CORRELATION ANALYSIS OF TRANSIENT RESPONSE

Now that we have suggested some potentially useful methods for making acoustical measurements, it would be wise to explore briefly the mathematics of this method.

It can be shown directly that the correct application of correlation measuring techniques to bounded sound fields will yield the impulse response of that bounded sound field[5].

Referring to Fig. 3.1, we assume that the noise source is delivering a signal having the complex spectrum $F_s(\omega)$ to the two channels of the correlator. If the networks have transfer functions $H_1(S)$ and $H_2(S)$ respectively, then the output of the networks are:

$$F_1(\omega) = H_1(j\omega)\ F_s(\omega) \qquad (3.2\text{-}1)$$

$$F_2(\omega) = H_2(j\omega)\ F_s(\omega) \qquad (3.2\text{-}2)$$

The cross power spectrum for $F_1(\omega)$ and $F_2(\omega)$ is

$$\Phi_{21}(\omega) = \lim_{T\to\infty} \frac{\pi}{T} \left[H_1(j\omega)\ F_s(\omega)\ H_2(-j\omega)\ \overline{F_s(\omega)}\right]$$

$$= \Phi_{ss}(\omega)\ H_1(j\omega)\ H_2(-j\omega) \qquad (3.2\text{-}3)$$

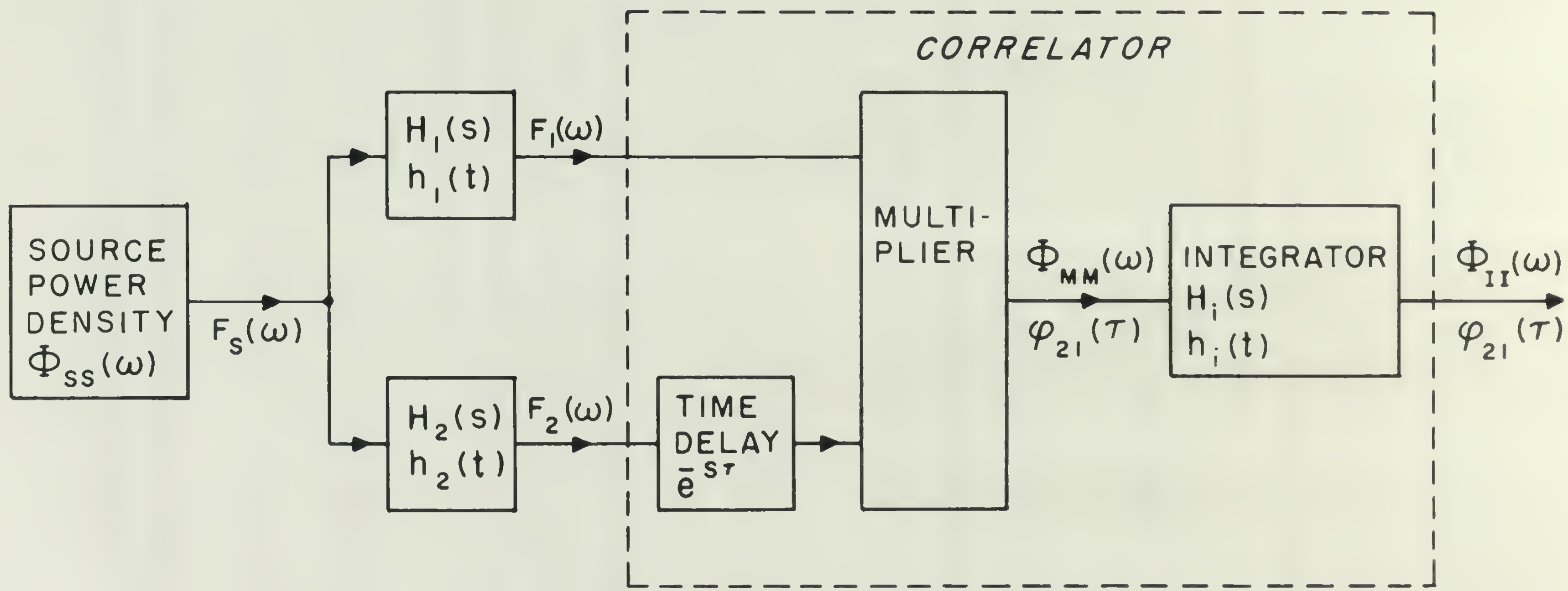


Fig. 3.1 Block diagram of correlator for analysis in the frequency domain.

where

$$\Phi_{ss}(\omega) \triangleq \lim_{T\to\infty} \frac{\pi}{T} \left|F_s(\omega)\right|^2 \qquad (3.2\text{-}4)$$

is the power density spectrum of the noise source.

Now the correlation is the Fourier cosine transform of the power density spectrum

$$\phi_{21}(\tau) = \int_{-\infty}^{\infty} \Phi_{ss}(\omega)\, H_1(j\omega)\, H_2(-j\omega)\, e^{j\omega\tau}\, d\omega \qquad (3.2\text{-}5)$$

From Fig. 3.1 we can see that frequency variation in the source can be considered as common terms to both channels one and two. This factor then can be considered as a constant, i.e.

$$\phi_{21}(\tau) = \Phi_{ss} \int_{-\infty}^{\infty} H_1(j\omega)\, H_2(-j\omega)\, e^{j\omega\tau}\, d\omega \qquad (3.2\text{-}6)$$

If we now consider the special case where $H_2 = 1$ and we wish to auto-correlate the output and input of the system having a transfer function of $H_1(j\omega)$, the cross-correlation function is

$$\phi_{21}(\tau) = \Phi_{ss} \int_{-\infty}^{\infty} H_1(j\omega)\, e^{j\omega\tau}\, d\omega =$$

$$2\pi\, \Phi_{ss}\, h_1(t) \qquad (3.2\text{-}7)$$

Thus, the cross-correlation function is exactly proportional to the impulse response of the system having

the transfer function $H_1(j\omega)$. If the channel having the transfer function $H_1(j\omega)$ is a room, then the correlation function is proportional to the impulse response of the room.

We have talked a little about the nature of the correlation function; let us now consider its shape for a correlation analysis of the sound field of a room. For this purpose, we shall first define a few terms.

Using Laplace Transformation we made the definition

$$F_1(\omega) \triangleq \frac{1}{2\pi} \int_{-\infty}^{\infty} f_1(t)\, e^{-j\omega t}\, dt \qquad (3.2\text{-}8)$$

whence

$$f_1(t) = \int_{-\infty}^{\infty} F_1(\omega)\, e^{j\omega t}\, d\omega$$

The Multiplication Theorem states:

$$2\pi \int_{-\infty}^{\infty} F_1(\omega)\, F_2(\omega)\, e^{j\omega\tau}\, d\omega = \int_{-\infty}^{\infty} f_1(t)\, f_2(t-\tau)\, dt \qquad (3.2\text{-}9)$$

Now dividing each side by $2T$ where T = time, and then taking the limit as $T \to \infty$

$$\lim_{T\to\infty} \frac{2\pi}{2T} \int_{-\infty}^{\infty} F_1(\omega)\, \overline{F_2(\omega)}\, e^{j\omega t}\, d\omega =$$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f_1(t)\, f_2(t-\tau)\, dt \triangleq \varphi_{21}(\tau) \qquad (3.2\text{-}10)$$

In the application of correlation techniques to the transient response of rooms $f_1(t)$ is the voltage produced at the loudspeaker terminals by a wide load noise-voltage generator and $f_2(t)$ is the voltage output of a microphone placed at some observation place in the sound field.

The voltage produced by the microphone is $f_2(t)$. It consists of the sum of $f_1(t)$ modified by the transfer characteristic H of the loudspeaker and microphone, plus a summation of reflections each of which equals $f_1(t)$ delayed by a time equal to the path length of transmission, l, divided by the speed of sound c; and modified by H and a transfer characteristic B associated with its transmission path.

The frequency domain representation of such a voltage is

$$F_2 = F_1 H B_o e^{j\omega l_o/c} + F_1 H B_1 e^{j\omega l_1/c} +$$

$$F_1 H B_2 e^{j\omega l_2/c} + - - - \qquad (3.2\text{-}11)$$

The B's are complex coefficients which account for the phase shifts at the reflecting surfaces, the attenuation in amplitude associated with inverse square dimunition of radiated sound intensity, and the attenuation in amplitude occurring at the reflecting surface.

$$B \triangleq |B| \; e^{j\omega\beta}$$

therefore

$$F_2 = \sum_{n=1}^{\infty} F_s H e^{-j\omega l_n/c} B_n \qquad (3.2\text{-}12)$$

$$= F_s H \sum_{n=1}^{\infty} |B_n| \quad e^{j(\beta_n - \omega l_n/c)}$$

If we now make the definition:

$$S = \text{power density spectrum} = \lim_{T\to\infty} \frac{\pi |F(\omega)|^2}{T} \qquad (3.2\text{-}13)$$

then it follows that

$$F(\omega) = \lim_{T\to\infty} \sqrt{\frac{T.S}{\pi}} \quad e^{j\phi} \qquad (3.2\text{-}14)$$

where

$$F(\omega) \triangleq |F(\omega)| \quad e^{j\phi}$$

Now substituting in the multiplication theorem:

$$\phi_{12}(\tau) = \lim_{T\to\infty} \int_{-\infty}^{\infty} \frac{\pi}{T} \sqrt{\frac{T S_1}{\pi} \quad \frac{T S_2}{\pi}}$$

$$e^{-j(\phi_1 - \phi_2 + \omega t)} \, d\omega$$

$$= \int_{-\infty}^{\infty} \sqrt{S_1 S_2} \quad e^{j(\omega t + \phi_2 - \phi_1)} \, d\omega \qquad (3.2\text{-}15)$$

When we correlate F_1 and F_2 using equation 3.2-15 and 3.2-12

$$\phi_{12}(\tau) = \int_{-\infty}^{\infty} H \sqrt{S_1 S_1} \quad e^{j\omega\tau} \left[\sum_{n=1}^{\infty} |B_n| \right.$$

$$\left. e^{j(\beta_n - \omega l_n/c)} \right] d\omega$$

[illegible]

[illegible]

$$[illegible]$$

$$[illegible]$$

[illegible]

[illegible]

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

$$\varphi_{12}(\tau) = \int_{-\infty}^{\infty} H\, S_1 \left[\sum_{n=1}^{\infty} |B_n|\, e^{j(\beta_n + \omega(\tau - l_n/c))}\right] d\omega \tag{3.2-16}$$

Equation 3.2-16 then gives the cross-correlation $\varphi_{12}(\tau)$ between the input voltage to the loudspeaker and the output voltage of the microphone in terms of the power density spectrum (S) of a noise voltage generator, the combined transfer function (H) of the loudspeaker and microphone, and the complex transfer function of the transmission path. This latter transfer function $|B_n|\, e^{j\beta_n}$ describes the complex modification in the sound pressure amplitude spectrum caused by its traveling the distance l_n involving any number of reflections.

For the idealized case where: (a) source radiates noise of a flat spectrum from ω_1 to ω_2 and zero outside of this spectrum; (b) the transfer responses of the loudspeaker, microphones, and walls are $1\underline{/0^\circ}$; it has been shown[10] that the cross-correlation curve for the loudspeaker and microphone voltages will have equal peaks for each point where the time delay corresponds to the transit time from the source to the microphone and that the cross-correlation curve will damp out to a small value between these points. The rapidity with which it damps out will depend upon the bandwidth $(\omega_2 - \omega_1)$ and the term $(\omega_2 + \omega_1)$.

With the foregoing mathematics in mind we may now visualize a room model that will be useful for correlation analysis. This model will be exactly the same as the model mentioned previously in section 3.1 in connection with pulse analysis of rooms. In this model the walls are replaced by an array of image sources each of which is located at a distance equal to the length of the transmission path of the reflected sound. The output of each source may be modified in accordance with the number of reflections that occur in the path of transmission from the source to the observer's position.

In their discussion of pulse statistics theory, Bolt, Doak, and Westervelt[8] developed the mathematics of such a model. In their model each image source occupies a separate image replica of the physical room. These image rooms or cells are designated by three numbers (l, m, n) each of which can take all integral values from minus infinity to plus infinity. The notation designating the image sources is also an indicator of the number of reflections that the signal undergoes in traveling from the source to the observer.

We shall now make a few simple adaptations of this spatial model to aid us in visualizing the correlation analysis of the problem. All of the sources shall be

grouped in accordance with the number of reflections that occur in transit from the source to the observer. Thus the sources whose sound is reflected twice in transit are designated as second order image sources.

The number of image sources in each order depends upon the room geometry. For example, the model for a rectangular room can be broken down into its image components as follows:

SOURCES	NUMBER
Actual	1
1st Order Image	6
2nd Order Image	18
3rd Order Image	38
etc.	

An outline of the values enclosed by the actual plus the six first order image sources for a rectangular room is shown in Fig. 3.2. The number of image sources of each order can be computed from the formula

$$\text{Number} = 2 + 4N + 8 \sum_{k=0}^{N-1} k \qquad (3.3\text{-}1)$$

where N is the order of the reflection and $k = (N-1)$. For example, for the fourth order array, there will be $2 + 16 + 8\,(3+2+1) = 66$ image sources.

Other factors of importance in correlation analysis are the distance between image sources, and the distance

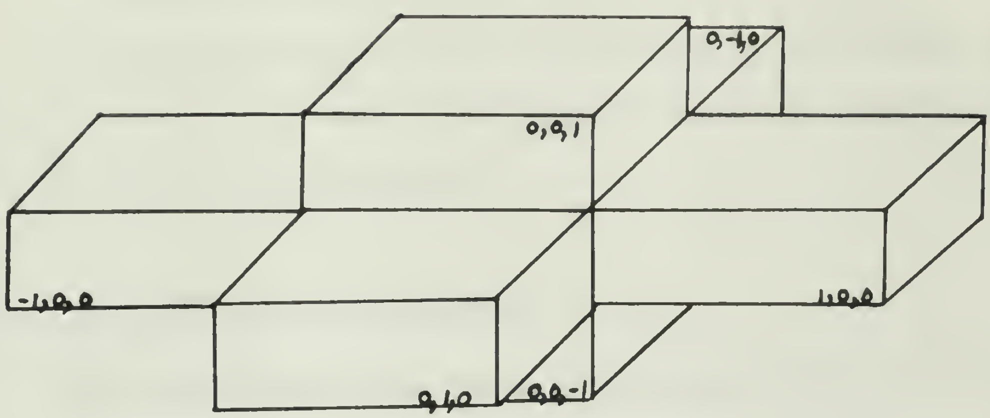


Fig. 3.2 Volume enclosed by actual and first order image sources for a rectangular room.

from the image sources to the observer's position. These factors may be computed from formulas 9 and 11 of reference 8.

The advantages gained by reference to this spatial model of a room are threefold:

a) the system gives a visual presentation of the problems encountered in space or time filtering.

b) the results of the correlation analysis may be compared with those of other researchers who used pulse methods of analysis.

c) the mathematics for pulse analysis developed by the use of this model may be applied directly to correlation analysis.

3.3 EXPERIMENTAL MEASUREMENTS

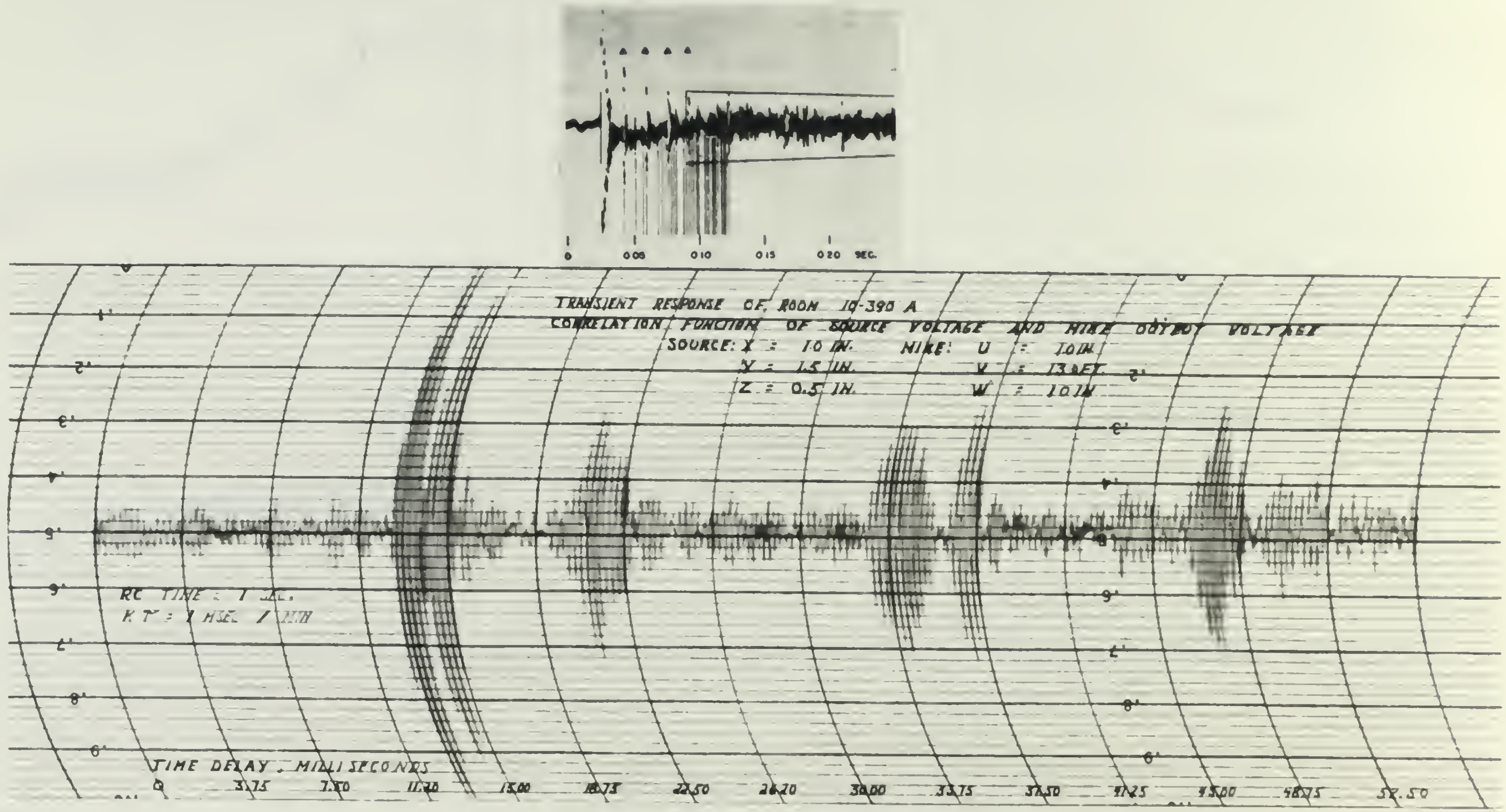
The experimental measurements of the transient response of rooms were taken by cross correlating between the input voltage to a loudspeaker and the output of a microphone located in the room. A hard plaster room, Room 10-390-A, which had been previously analyzed by pulse methods[8] was selected for the preliminary tests. Both the source and microphone, as well as the test positions in the room, were selected to give results comparable to those obtained by Bolt, Doak, and Westervelt[8].

The results for one particular test run are shown in Fig. 3.3. Shown also are the comparable results of a pulse analysis. In this run, the source, a Model 175 James Lansing horn driving unit with a Model 1217-1290 James Lansing throat and lens assembly, was placed in the corner of the room facing the wall. The microphone, an Altec 21-BR-150A was located 13 feet from the source along the short wall of the room.

The pertinent dimensions are:

Room dimensions:	length	L_x = 23 feet
	width	L_y = 13.4 feet
	height	L_z = 8.4 feet
Source Position:		x = 1.0 inches
		y = 1.5 inches
		z = 0.5 inches
Mike Position:		u = 1.0 inches
		v = 13 feet
		w = 1.0 inches

The test signal was random noise generated by an electronic noise voltage generator and amplified by a Model 20-W2 McIntosh amplifier. A 1/3 octave band analysis of the loudspeaker voltage and the microphone voltage is shown in Fig. 3.4. The loudspeaker voltage and the microphone voltage were recorded simultaneously on separate tracks of a twin track tape recorder. In reducing the data, a section of the recording tape was



0 0.05 0.10 0.15 0.20 SEC.
TRANSIENT RESPONSE OF ROOM 10-390 A
CORRELATION FUNCTION OF SOURCE VOLTAGE AND MIKE OUTPUT VOLTAGE
SOURCE: X = 1.0 IN. MIKE: U = 10 IN.
Y = 1.5 IN. V = 13.0 FT.
Z = 0.5 IN. W = 10 IN.
RC TIME = 1 SEC.
TIME DELAY, MILLISECONDS
0 3.75 7.50 11.20 15.00 18.75 22.50 26.20 30.00 33.75 37.50 41.25 45.00 48.75 52.50

Figure 3.3

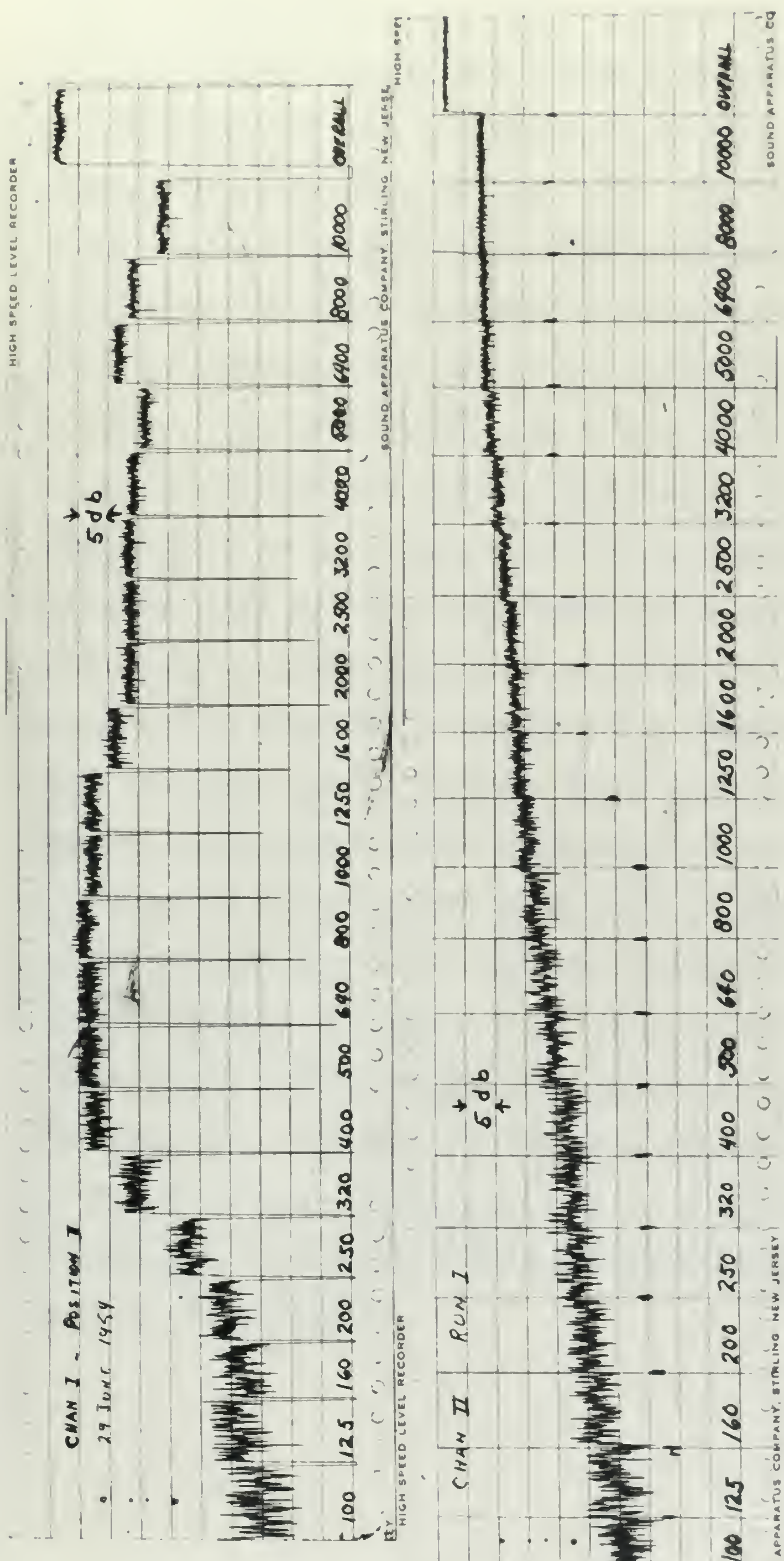

HIGH SPEED LEVEL RECORDER
CHAN I - POSITION I
29 JUNE 1954
5 db
100 125 160 200 250 320 400 500 640 800 1000 1250 1600 2000 2500 3200 4000 5000 6400 8000 10000 OVERALL
SOUND APPARATUS COMPANY, STIRLING NEW JERSEY
CHAN II
RUN I
5 db
100 125 160 200 250 320 400 500 640 800 1000 1250 1600 2000 2500 3200 4000 5000 6400 8000 10000 OVERALL

Figure 3.4

cut and made into a tape loop as shown in Fig. 2.5. The microphone voltage was filtered so that its spectrum was that passed by a 1/3 octave band filter with a geometric mean frequency of 4000 cps. The filtered microphone voltage was fed into channel one of the correlator time delay and the unfiltered loudspeaker voltage was placed across channel two of the correlator. An RC integrator time of 1 second and a time delay scanning rate of 1 millisecond per minute were used. The resulting linear plot of the cross correlation function together with the plot for the comparable pulse test is shown in Fig. 3.3. In this figure, each group of peaks in the correlation function represents the arrival of the sound from an image source in much the same way as the peaks in the accompanying pulse analysis. An accompanying log plot of the correlation was made. This plot is not presented because of the extreme physical length of the plotting tape. However, Fig. 3.5 contains a plot of the peak values of the correlation curve as a function of C for Fig. 3.2. Plotted on the same coordinates is a graph of the peak pulse pressure level as a function of distance traveled. The points for this second curve are taken from Fig. 4 (reference 8).

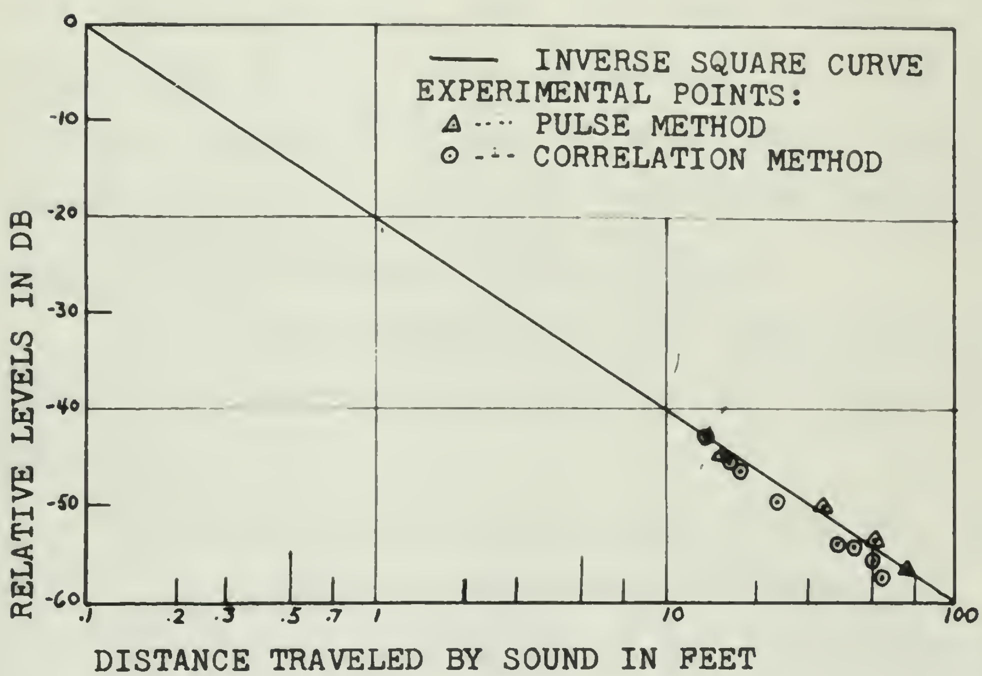


Fig. 3.5 Plot of the transient response of Room 10-390A for pulse and correlation analysis.

3.4 DISCUSSION OF EXPERIMENTAL RESULTS

The data on pulse statistics analysis[8] shows the spacing of pulses arriving at the microphone position when both source and receiver are in the corner of the room. The correlation analysis of the transient response of the room should show comparable peaks in the correlation curve for time delays corresponding to these pulse spacings. The correlation curve, Fig. 3.3, shows discrete peaks at delay times of 12.7, 13.6, 21.2, 34.4, 37.2, and 47.5 milliseconds. Reference to Fig. 14 of reference 8 shows that these delay times correspond roughly with the transit times for the various reflected pulses.

The relative amplitudes of the peaks of the correlation give a quantitative measure of the combined effects of:

1. Attenuation of signal caused by inverse square diminuation of radiated sound intensity.
2. Attenuation of signal caused and decrease in coherence of radiation from the various image sources caused by
   a) The transfer function of one reflecting surface for the case of non-degenerate discrete first order image sources.
   b) The combined transfer function of all the reflecting surfaces for higher order non-degenerate image sources.

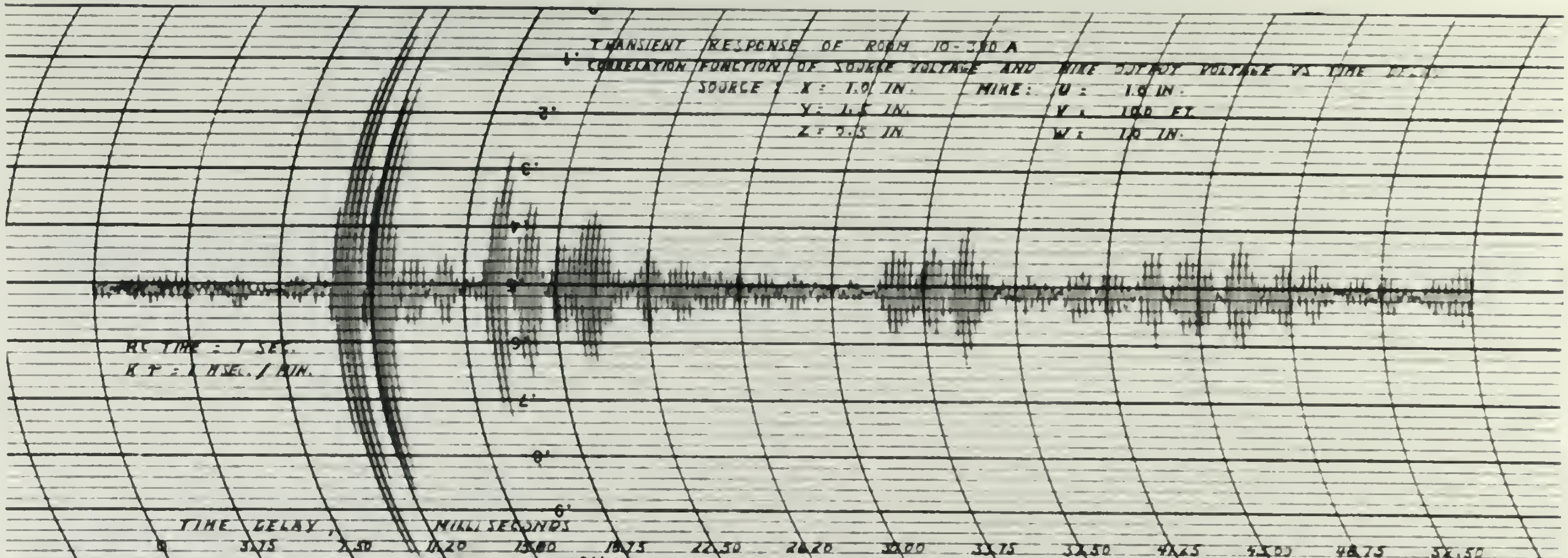
c) Decrease in coherence of the radiation from groups of degenerate image sources.

Thus, the plot of the correlation function vs time delay is a measure of the short term transient response of the room for a particular source and receiver position.

Using correlation analysis, it is also possible to derive a value for reverberation time based on the transient response of the room. The reverberation time simply corresponds to the difference in delay time between the peak in the curve corresponding to the arrival of sound from source 0, 0, 0 and a peak in the correlation curve which is one-millionth of, or 60 db down from the peak of the direct sound. This time may be found by extrapolating the rate of decay indicated by the short term transient response. For the example discussed in the previous section where both source and receiver were in the corner, the correlation curve decayed 20 db in the interval τ = 10 milliseconds to τ = 67 milliseconds. The reverberation time based on this rate of decay is 3.14 seconds. This value, then, is a measure of the reverberation time for a 1/3 octave band signal having a geometric mean frequency of 4000 cps.

In line with the experimental pulse analysis of reference 8, additional transient response data corresponding to the data of Figs. 7 through 16 has been collected and reduced by means of the analog correlator.

This additional data with explanatory notes is presented in Figs. 3.6, 3.7, 3.8, and 3.9. For the convenience of the reader, a copy of reference 8 is included in the Appendix.



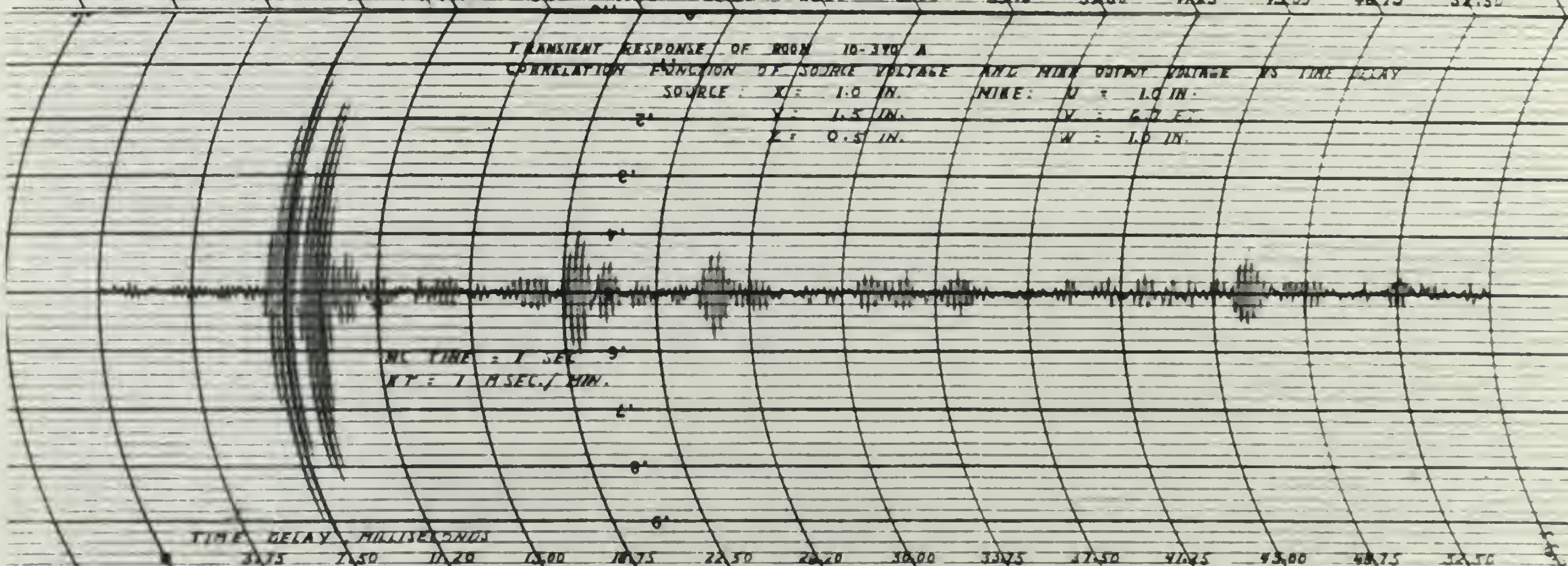


Figure 3.6 Top
Figure 3.7 Bottom

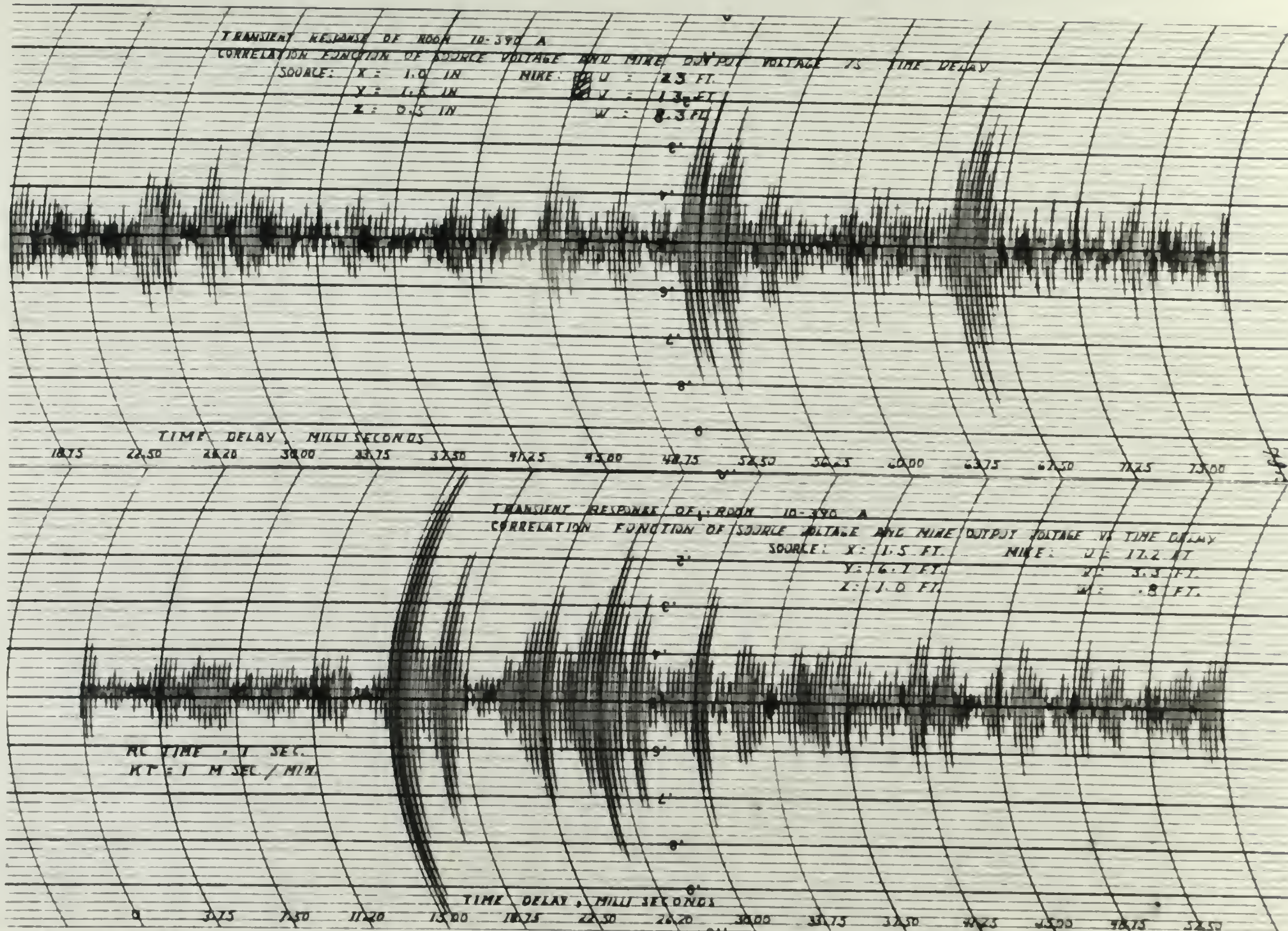


Figure 3.8 Top
Figure 3.9 Bottom

IV. MEASUREMENTS OF DIFFUSION

4.1 INTRODUCTION

In the literature on room acoustics, much attention has been directed toward defining properties of sound fields which seem to have a perceptible and quantitative effect upon the goodness of the acoustical environment. One such property is "diffusion." Without delving into the technical, psycho-acoustic aspects of the problem, it would be advisable to discuss some of the more intuitively perceptible aspects of acoustics which are attendant to the property of diffusion.

A sound field is defined as being completely diffuse if it has uniform energy density within the region considered, and if the direction of propogation of the wave fronts at any arbitrary selected points are wholly random in distribution. In discussing some of the trends in the study of sound waves in rooms[11] Morse and Bolt say, "It now appears that reverberation time alone is not always a sufficient measure of auditorium excellence. It is desirable to have a mean square pressure as nearly uniform as possible over the seating area. It is also important to have at least a certain percentage of the sound reach the listener directly from the speaker, and less than a certain percentage reach the listeners indirectly, after reflection from any single surface; ... " In the light

of the definition of diffuseness, the above quotation prescribes that for any sound field, there is an optimum degree of diffusion which produces a good acoustical environment. One may reason that this optimum degree of diffusion will depend upon the use to which the sound field is to be adopted i.e. lecture hall, broadcast studio, cinema, music hall; and more particularly, upon the acoustical tradition of the population in the sound field. For instance, the population of Britain may be said to have slightly different preferences than the population of the United States. These preferences would probably be based upon the way the British have been used to hearing their music or lectures and also upon the peculiarities of the language. The degree of diffusion in any sound field will in general be a function of the location of the source, the shape of the boundaries enclosing the sound field, and the shape, absorption coefficient and distribution of the surfaces in the sound field.

As a sort of intuitive example of the effects of diffusion let us visualize a full symphony orchestra playing in an open field. Generally, the conductor would be able to hear clearly only those instruments which were near to him or whose radiation was directive and beamed in his direction. The musicians would be able to hear proportionally less of the total effect being produced by the orchestra as a group.

Under these circumstances, it is probable that a very poor performance would be the result of this lack of diffusion or mixing. On the other hand, if say a small string ensemble was performing in a very diffuse broadcast studio, each member of the ensemble would be able to hear the combined effect of the group, but he would hear relatively little of the sound that his particular instrument was producing. Such a condition would again result in an inferior performance. With the proper degree of diffusion, however, it is theorized that conditions would be optimum for the production of the desired acoustical effect.

In a technical report Number B 058, Serial No. 1953/29, dated October 1953, the British Broadcasting Company discussed the property of diffusion of sound fields. It was said that imperfect diffusion may appear in at least two measurable forms, (1) as a change in pressure with a change in position, the frequency of the sound remaining constant; (2) as a change in pressure with changes in frequency, the positions of source and receiver remaining unchanged. It was further stated that imperfect diffusion indicated in the steady state measurement by variation of pressure with frequency or position is shown by deviation in the pressure, time relations from an exponential decay curve. These statements are perfectly reasonable; however, the functional relationship between the degree of diffusion

and the transient response of a sound field is not apparent.

Correlation techniques seem to offer a method of obtaining a measure of the diffusion of a sound field which depends directly upon the definition of diffusion. Moreover, this method employs almost the same equipment and arrangements as are used for the transient response measurements.

4.2 CORRELATION METHOD OF DIFFUSION MEASUREMENT

Let us visualize the following experiment.

We shall place a microphone in the sound field created by a loudspeaker driven by a wide band noise voltage generator.

Now let us auto-correlate the output voltage of the microphone. If the noise voltage has an infinitely wide band flat spectrum, then it has been shown[5] that the auto-correlation function of the microphone voltage is exactly the combined impulse response of the loudspeaker and the microphone. For noise voltages of finite bandwidth, this auto-correlation function is the transient response of the system to a pulse having the corresponding spectral composition. In general, this auto-correlation function has a maximum peak for zero time delay and a first axis crossing at a time delay dependent upon the bandwidth of the

noise voltage $\omega_2 - \omega_1$ and the quantity $\omega_2 + \omega_1$. A series of such auto-correlations is shown in Fig. 4.1.

As explained in Chapter I, the bandwidth and the first axis crossing of the auto-correlation function vary inversely. That is, the wider the bandwidth, the shorter the difference in delay times between the peak of the correlation function and the first axis crossing. It is possible to measure the time delay corresponding to this first axis crossing with a high degree of accuracy. The time delay register associated with the analog correlation computer permits readings of time delay to the tenth of a millisecond. It is possible to interpolate between these register readings by the following method.

a) Into each channel of the time delay, feed a standard 10 kcps signal voltage.

b) Measure the phase difference of the outputs of the time delay for this 10 kcps signal.

c) Since the period of the 10 kcps signal is precisely 1/10 of a millisecond, it is possible to hand adjust the time delay mechanism to a value corresponding to any fraction of a period in terms of degrees of phase difference between the time delay outputs.

d) With the time delay mechanism thus adjusted, one can now remove the 10 kcps signal, apply the noise voltage, and measure the voltage output corresponding to the value of the auto-correlation curve for

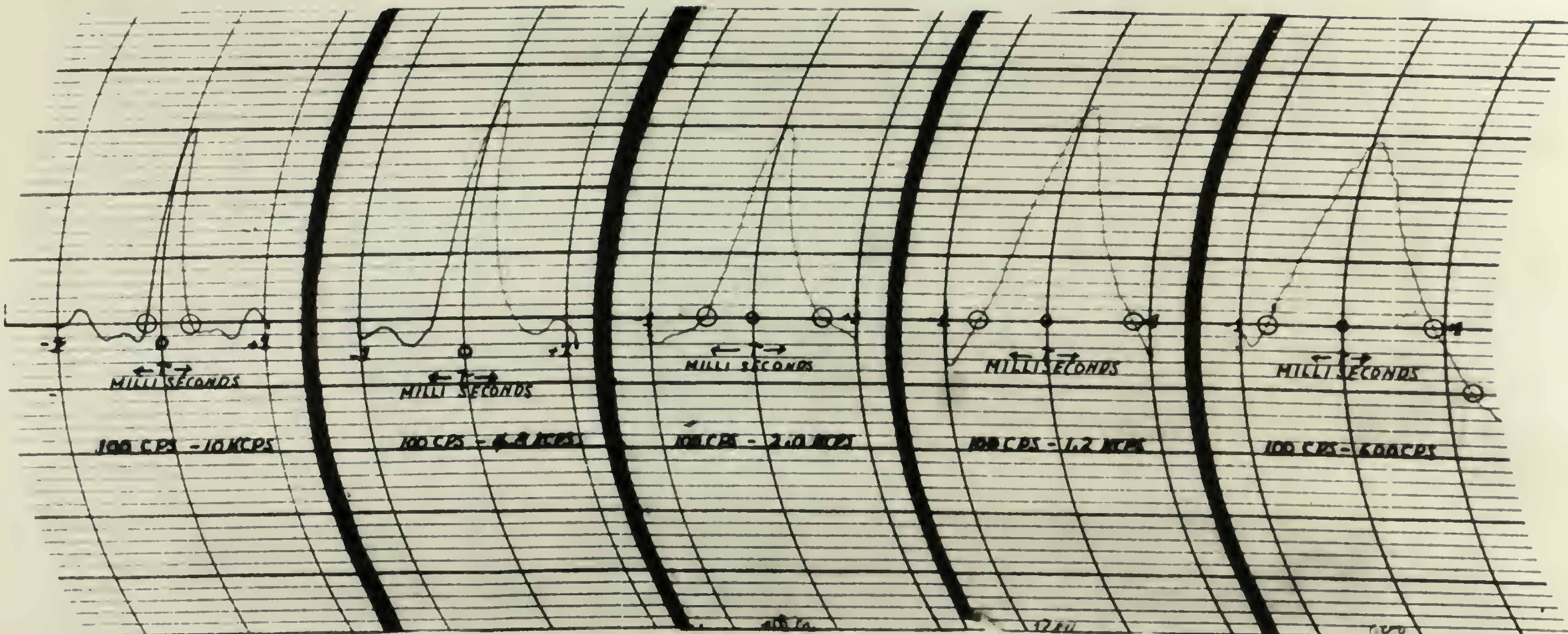


Figure 4.1 Plot of auto-correlation function for loudspeaker, microphone system for various noise voltage spectra.

that time delay.

Now, knowing the time delay of the axis crossing, one can place two such identical microphones in the sound field of a source radiating spherical sound waves. If the microphones are placed a distance apart equal to the time delay to the first axis crossing of the auto-correlation curve, and then orientated in such a way that the sound fronts from the source strike both mikes at the same instant (Fig. 4.2); then the cross-correlation function at zero time delay for the output voltage of the two microphones will correspond to the value of the auto-correlation curve for $= 0$. However, if the source microphone array is turned 90 degrees in a horizontal plane so that the sound waves strike one microphone, is subjected to a physical delay corresponding to the axis crossing time, and then strikes the second microphone (Fig. 4.3), then the corresponding cross-correlation function will be zero. Thus it is possible to obtain a plot of cross-correlation vs the orientation the microphones array in degrees from 0 to 90°. If we specify that zero degrees corresponds to the orientation when both mikes are equidistant from the source, then we can convert our previously measured accurate plot of the auto-correlation curve vs time delay to a plot of cross-correlation function vs orientation in degrees by means of the formula $= d \sin \theta$ where d is the mike

Fig. 4.2 Photograph of microphone array oriented at 0 degrees in sound field of loudspeaker.

Fig. 4.2 Photograph of microphone array oriented at 0 degrees in sound field of loudspeaker.

[illegible]
[illegible]

Fig. 4.3 Photograph of microphone array oriented at 90 degrees in sound field of loudspeaker.

separation in feet, c is the speed of sound in feet per millisecond, and is time delay in milliseconds. These calculations have been plotted for a noise spectrum flat from 100 cps to 10 kcps, having a roll off of 6 db per octave beyond these frequencies (Fig. 4.4).

The microphone array together with the multiplying and integrating circuits of the analog correlator comprise an instrument which gives a measure of the properties involved in the definition of diffusion.

Suppose the microphone array was placed in a perfectly diffuse field. Then, no matter which way the array was turned, the value of the cross-correlation function would be the same, since by definition, a diffuse field is one in which the wave fronts appear to be coming uniformly from all directions. It is clear that for variations in diffusion, the related cross-correlation curve would have some characteristic shape.

4.3 EXPERIMENTAL RESULTS

EXPERIMENT 1

Two Altec 21-BR-150A microphones were placed two inches apart in anechoic space in the sound field produced by an Altec 728-B closed-box-baffled loudspeaker. The cross correlation of the clipped output voltage of the microphones was plotted as a function of orientation, Fig. 4.5, from 0 degrees, Fig. 4.2, to 90 degrees,

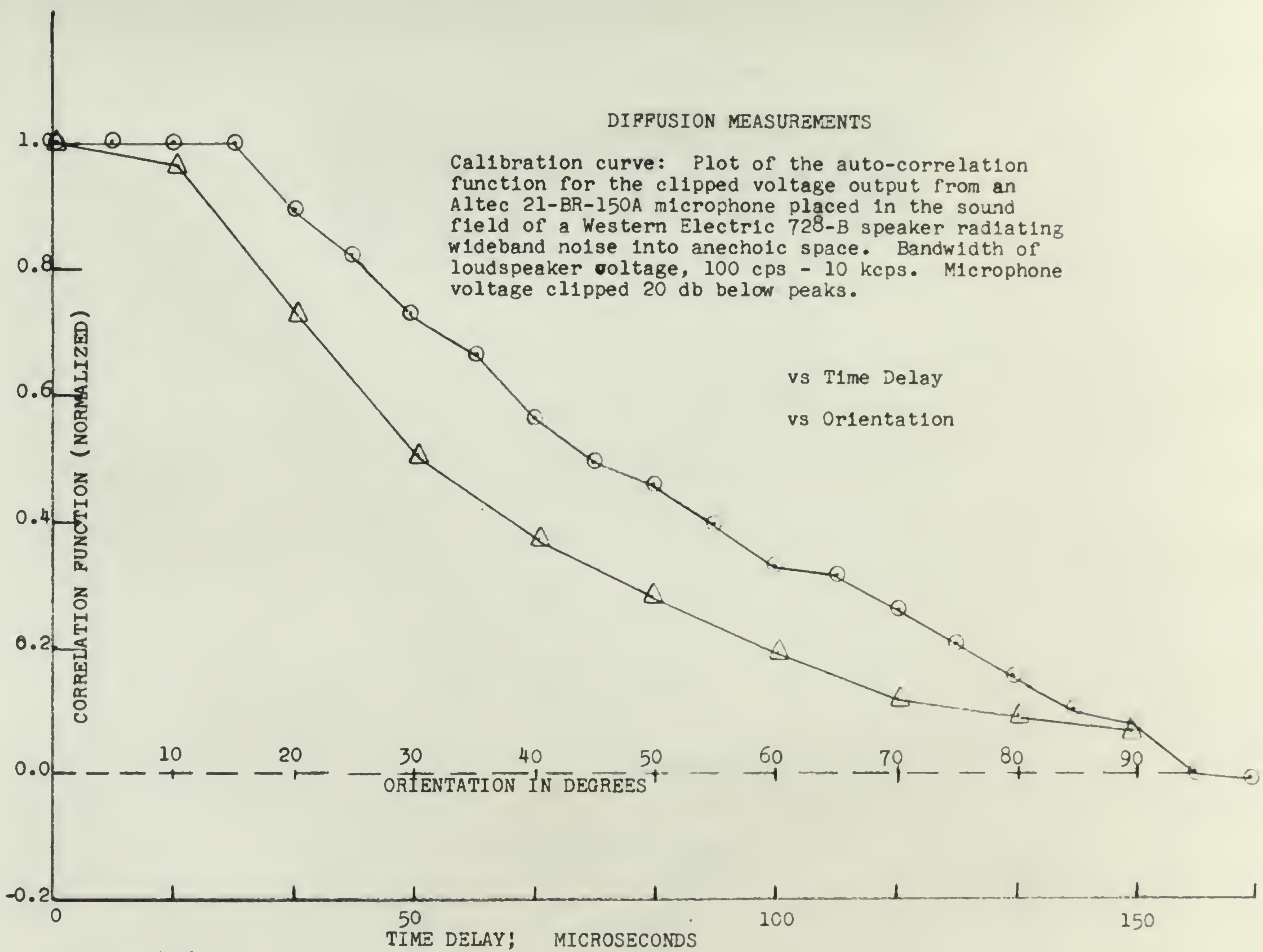

DIFFUSION MEASUREMENTS
Calibration curve: Plot of the auto-correlation function for the clipped voltage output from an Altec 21-BR-150A microphone placed in the sound field of a Western Electric 728-B speaker radiating wideband noise into anechoic space. Bandwidth of loudspeaker voltage, 100 cps - 10 kcps. Microphone voltage clipped 20 db below peaks.
vs Time Delay
vs Orientation
CORRELATION FUNCTION (NORMALIZED)
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
10
20
30
40
50
60
70
80
90
ORIENTATION IN DEGREES
0
50
100
150
TIME DELAY; MICROSECONDS

Figure 4.4

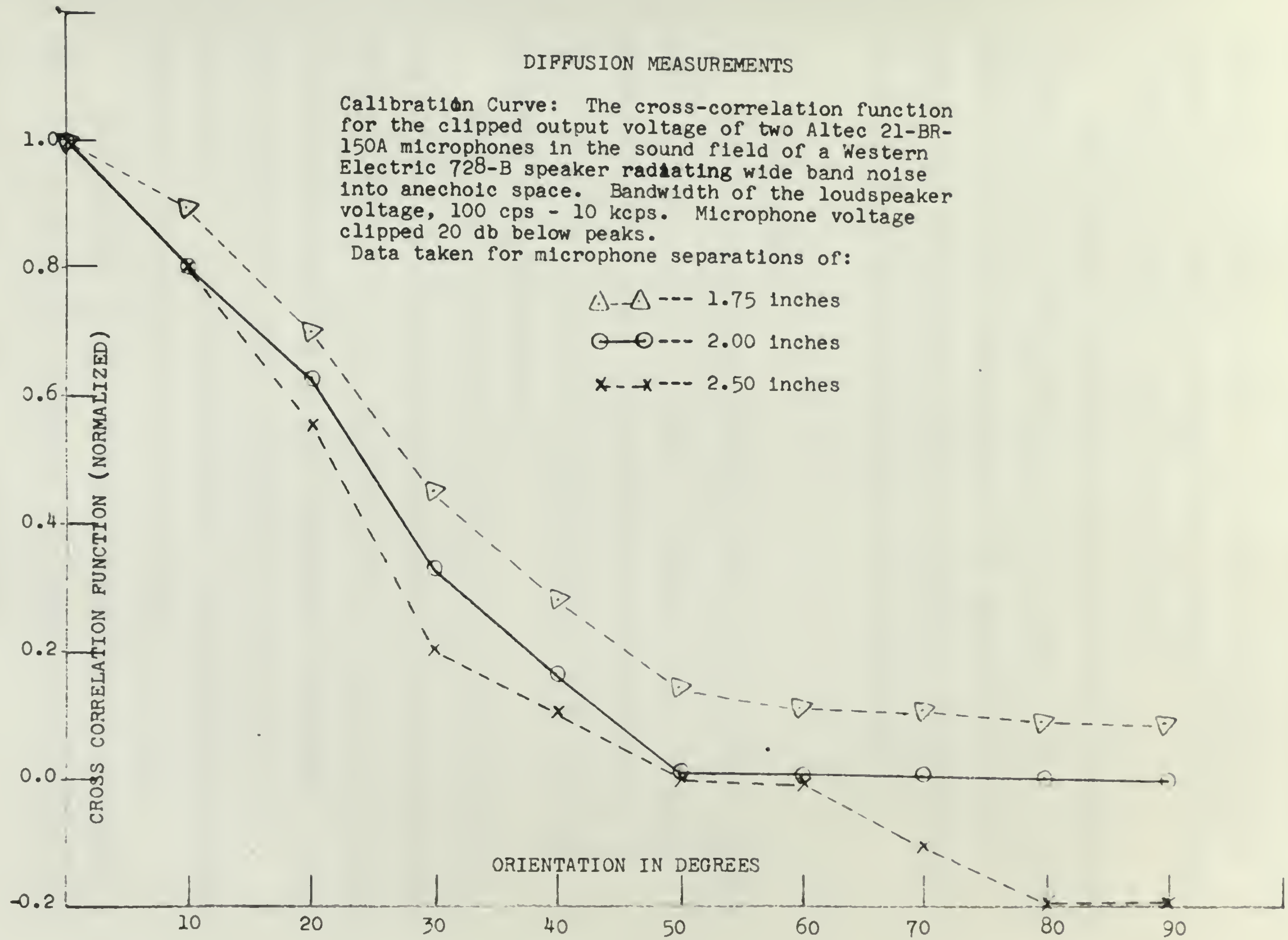

DIFFUSION MEASUREMENTS
Calibratiön Curve: The cross-correlation function for the clipped output voltage of two Altec 21-BR-150A microphones in the sound field of a Western Electric 728-B speaker radiating wide band noise into anechoic space. Bandwidth of the loudspeaker voltage, 100 cps - 10 kcps. Microphone voltage clipped 20 db below peaks.
Data taken for microphone separations of:
--- 1.75 inches
--- 2.00 inches
--- 2.50 inches
CROSS CORRELATION FUNCTION (NORMALIZED)
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
ORIENTATION IN DEGREES
10
20
30
40
50
60
70
80
90

Figure 4.5

Fig. 4.3. The curve in Fig. 4.5 is a measure of the diffusion of anechoic space. As should be expected, the source shows that the sound field produced by the loudspeaker in anechoic space is not diffuse.

The value of microphone separation used in this first experiment was obtained by trial and error. It is impossible to get the desired results by spacing the microphones the theoretically correct distance apart because of the finite size of the microphone diaphragms.

EXPERIMENT 2

The following experimental work was performed in the model studio, Room 20-F-009A, located in the Acoustics Laboratory at M.I.T. This studio was designed as a diffuse room. A physical description of this room is contained in reference 12.

With the loudspeaker facing the wall shown in Figs. 4.2 and 4.3, diffusion data was taken with the microphone array placed at various distances for a line perpendicular to the radiating face of the loudspeaker enclosure. The noise voltage and other equipment was the same as in Experiment 1. As in Experiment 1, the microphone output voltages were clipped 20 db below their rms value.

The results of the diffusion measurements together with the plot of the transient response of the studio are shown in Figs. 4.6 and 4.7 respectively.

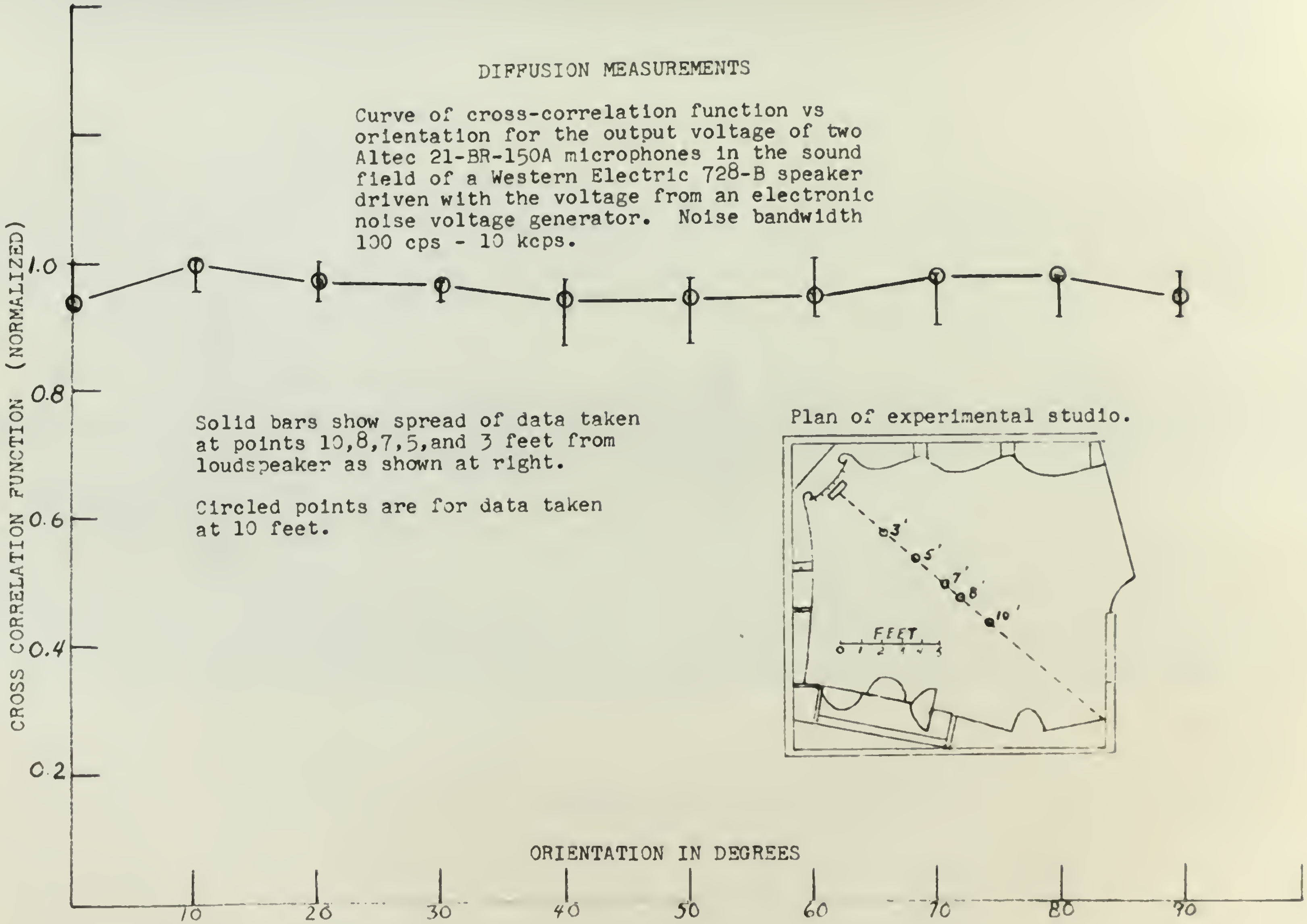

DIFFUSION MEASUREMENTS
Curve of cross-correlation function vs orientation for the output voltage of two Altec 21-BR-150A microphones in the sound field of a Western Electric 728-B speaker driven with the voltage from an electronic noise voltage generator. Noise bandwidth 100 cps - 10 kcps.
Solid bars show spread of data taken at points 10,8,7,5,and 3 feet from loudspeaker as shown at right.
Circled points are for data taken at 10 feet.
Plan of experimental studio.
FEET
0 1 2 3 4 5
3'
5'
7'
8'
10'
CROSS CORRELATION FUNCTION (NORMALIZED)
1.0
0.8
0.6
0.4
0.2
ORIENTATION IN DEGREES
10
20
30
40
50
60
70
80
90

Figure 4.6

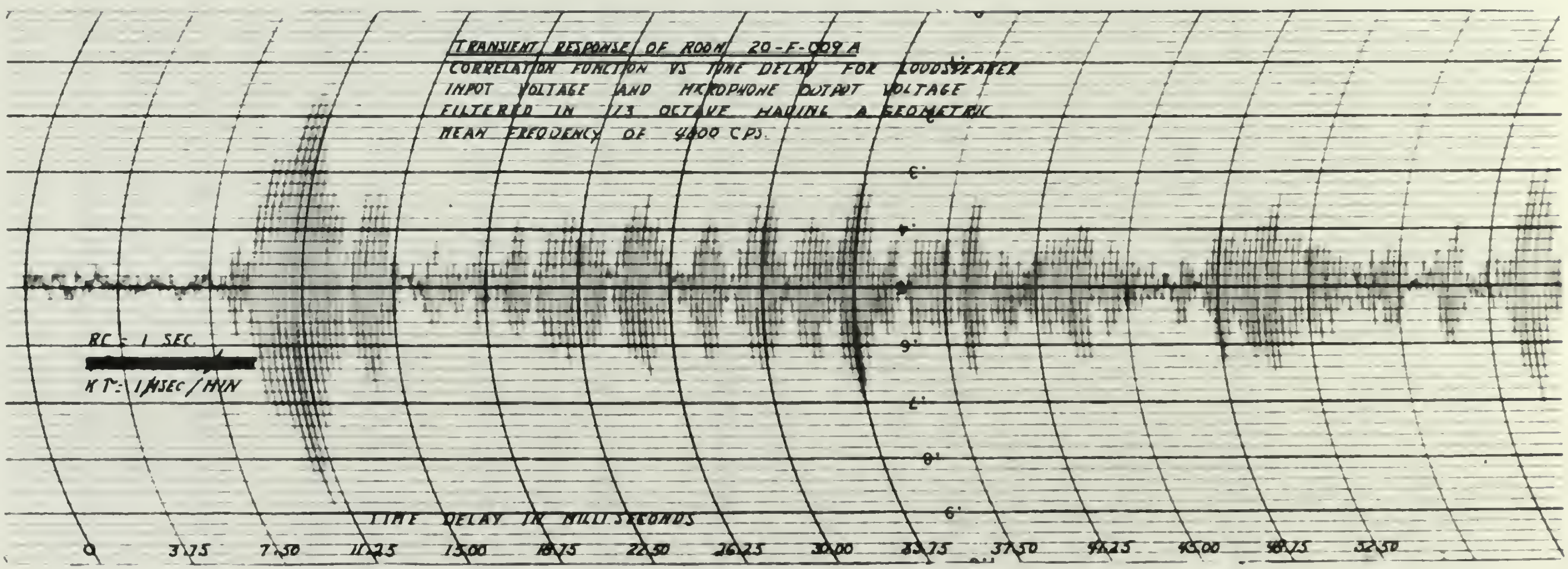

TRANSIENT RESPONSE OF ROOM 20-F-009A
CORRELATION FUNCTION VS TIME DELAY FOR LOUDSPEAKER
INPUT VOLTAGE AND MICROPHONE OUTPUT VOLTAGE
FILTERED IN 1/3 OCTAVE HAVING A GEOMETRIC
MEAN FREQUENCY OF 4800 CPS.
RC = 1 SEC.
KT = 1 MSEC/MIN
TIME DELAY IN MILLISECONDS
0
3.75
7.50
11.25
15.00
18.75
22.50
26.25
30.00
33.75
37.50
41.25
45.00
48.75
52.50

Figure 4.7

EXPERIMENT 3

The following experimental data was taken in Room 10-390-A, the hard plaster room. In these experiments the data was tape recorded and then brought into the laboratory and reduced.

The same equipment used in Experiments 1 and 2 was used in the following tests.

Diffusion data and transient response data was taken for various conditions of the room. For these tests, the loudspeaker was placed facing a corner of the room and the diffusion data was taken for various positions in the room. The positions were located at distances of 20, 15, 10, 8, and 5 feet from the loudspeaker on a line between the loudspeaker location and the diagonally opposite corner of the room. The transient response data was taken at the 15 foot position.

The above described data was taken, reduced, and plotted for the following room conditions:

a) All walls of the room were bare. Figs. 4.8 and 4.9.

b) One long wall of the room was covered with a highly absorbent Fiberglas filled quilt. Figs. 4.10 and 4.11.

c) One long wall and one short wall were covered with quilting. Figs. 4.12 and 4.13.

d) Two short walls and one long wall were covered with quilting. Figs. 4.14 and 4.15.

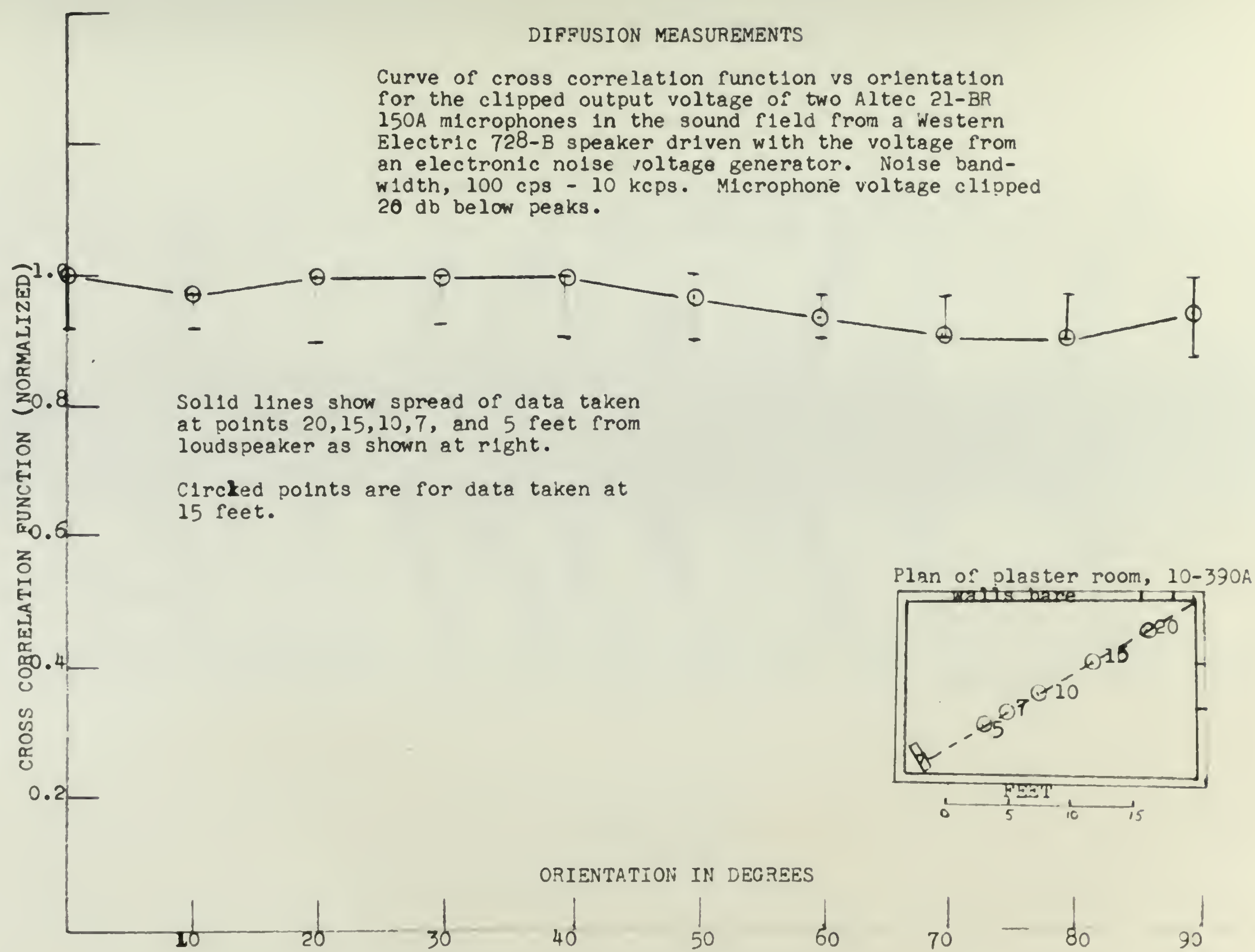

DIFFUSION MEASUREMENTS
Curve of cross correlation function vs orientation for the clipped output voltage of two Altec 21-BR 150A microphones in the sound field from a Western Electric 728-B speaker driven with the voltage from an electronic noise voltage generator. Noise bandwidth, 100 cps - 10 kcps. Microphone voltage clipped 20 db below peaks.
Solid lines show spread of data taken at points 20,15,10,7, and 5 feet from loudspeaker as shown at right.
Circled points are for data taken at 15 feet.
Plan of plaster room, 10-390A
walls bare
FEET
0
5
10
15
20
CROSS CORRELATION FUNCTION (NORMALIZED)
1.0
0.8
0.6
0.4
0.2
ORIENTATION IN DEGREES
10
20
30
40
50
60
70
80
90

Figure 4.8

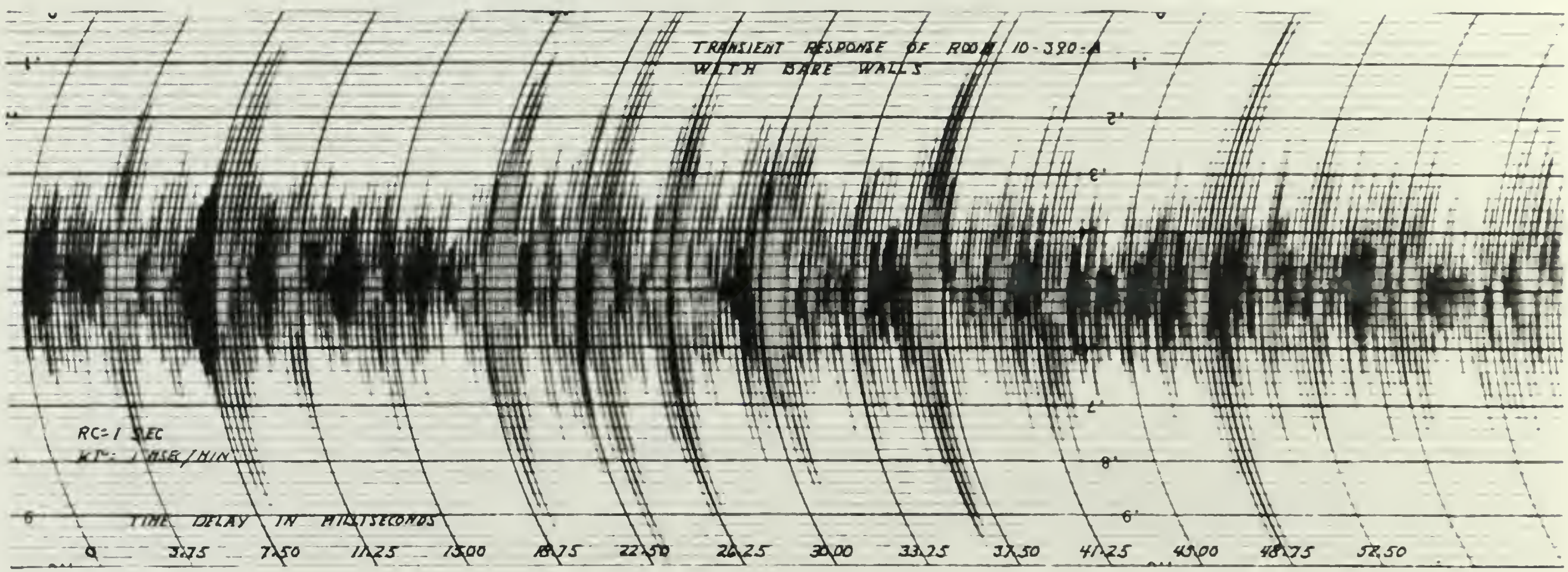

TRANSIENT RESPONSE OF ROOM 10-390-A
WITH BARE WALLS
RC=1 SEC
TIME DELAY IN MILLISECONDS
0
3.75
7.50
11.25
15.00
18.75
22.50
26.25
30.00
33.75
37.50
41.25
45.00
48.75
52.50

Figure 4.9

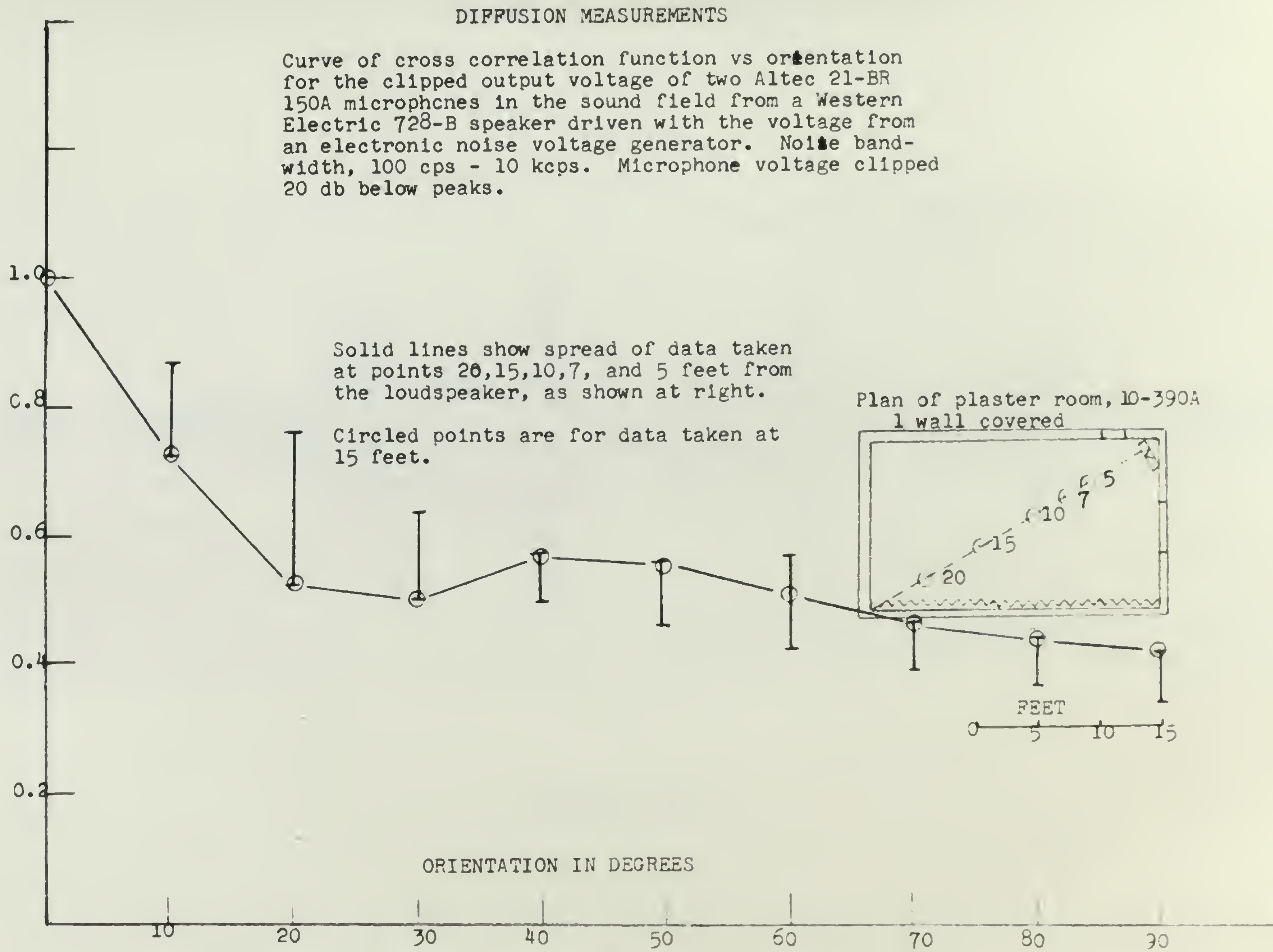

DIFFUSION MEASUREMENTS
Curve of cross correlation function vs orientation for the clipped output voltage of two Altec 21-BR 150A microphones in the sound field from a Western Electric 728-B speaker driven with the voltage from an electronic noise voltage generator. Noise band-width, 100 cps - 10 kcps. Microphone voltage clipped 20 db below peaks.
Solid lines show spread of data taken at points 20,15,10,7, and 5 feet from the loudspeaker, as shown at right.
Circled points are for data taken at 15 feet.
Plan of plaster room, 10-390A
1 wall covered
20
15
10
7
5
FEET
0
5
10
15
1.0
0.8
0.6
0.4
0.2
ORIENTATION IN DEGREES
10
20
30
40
50
60
70
80
90

Figure 4.10

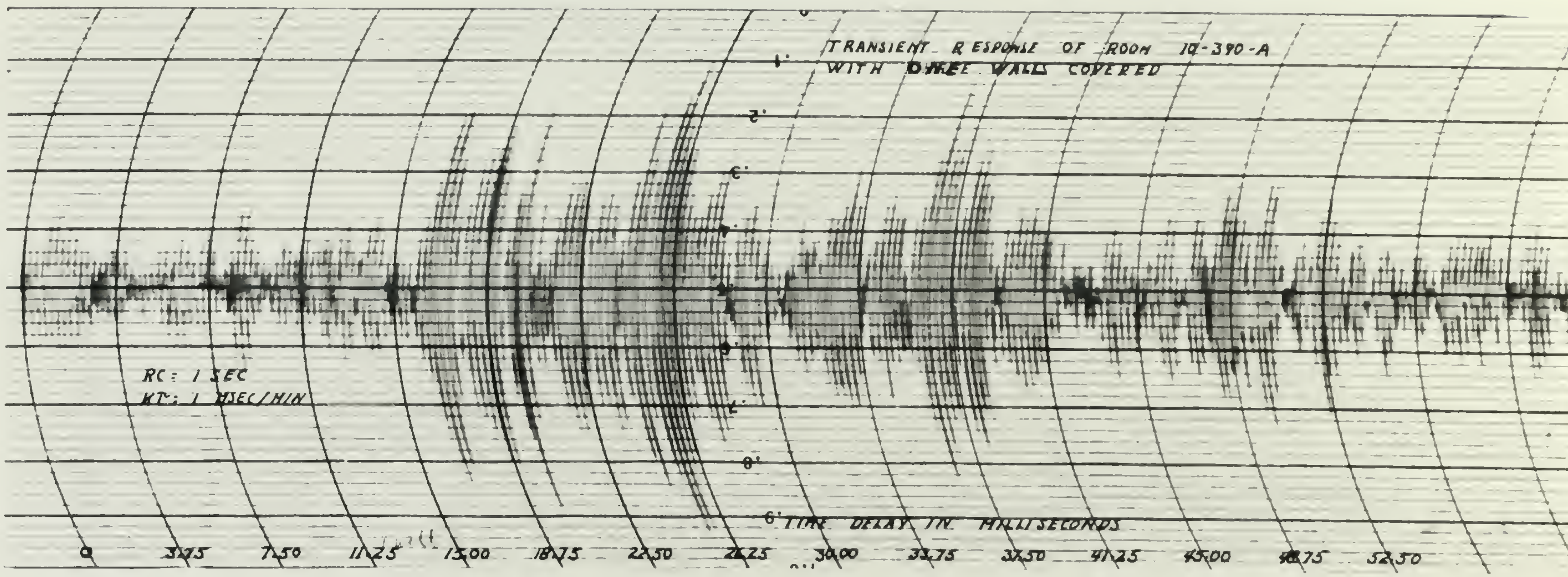

TRANSIENT RESPONSE OF ROOM 10-390-A
WITH THREE WALLS COVERED
RC = 1 SEC
KT: 1 MSEC/MIN
TIME DELAY IN MILLISECONDS
0
3.75
7.50
11.25
15.00
18.75
22.50
26.25
30.00
33.75
37.50
41.25
45.00
48.75
52.50

Figure 4.11

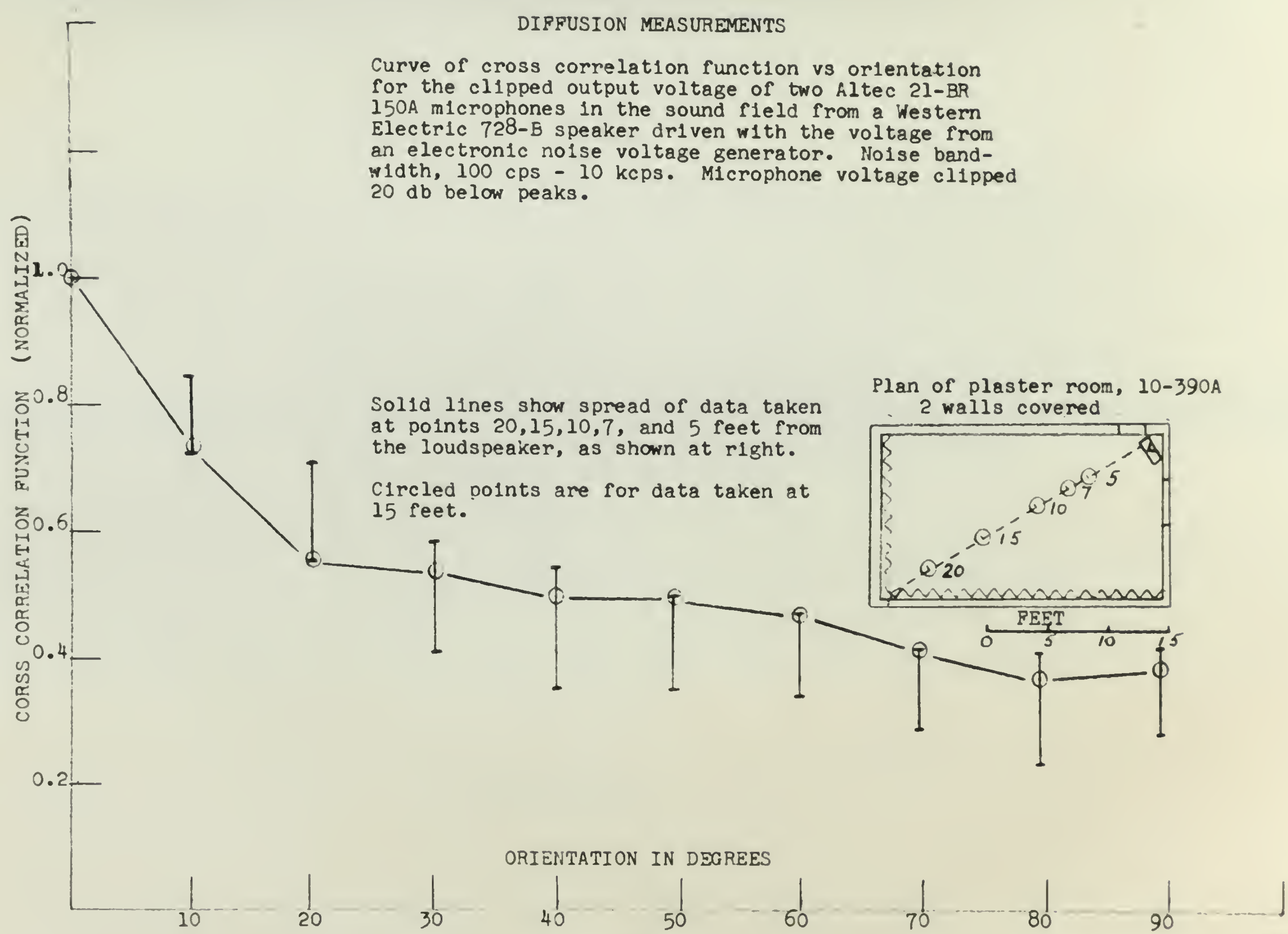

DIFFUSION MEASUREMENTS
Curve of cross correlation function vs orientation for the clipped output voltage of two Altec 21-BR 150A microphones in the sound field from a Western Electric 728-B speaker driven with the voltage from an electronic noise voltage generator. Noise bandwidth, 100 cps - 10 kcps. Microphone voltage clipped 20 db below peaks.
Solid lines show spread of data taken at points 20,15,10,7, and 5 feet from the loudspeaker, as shown at right.
Circled points are for data taken at 15 feet.
Plan of plaster room, 10-390A
2 walls covered
FEET
0
5
10
15
20
CORSS CORRELATION FUNCTION (NORMALIZED)
1.0
0.8
0.6
0.4
0.2
ORIENTATION IN DEGREES
10
20
30
40
50
60
70
80
90

Figure 4.12

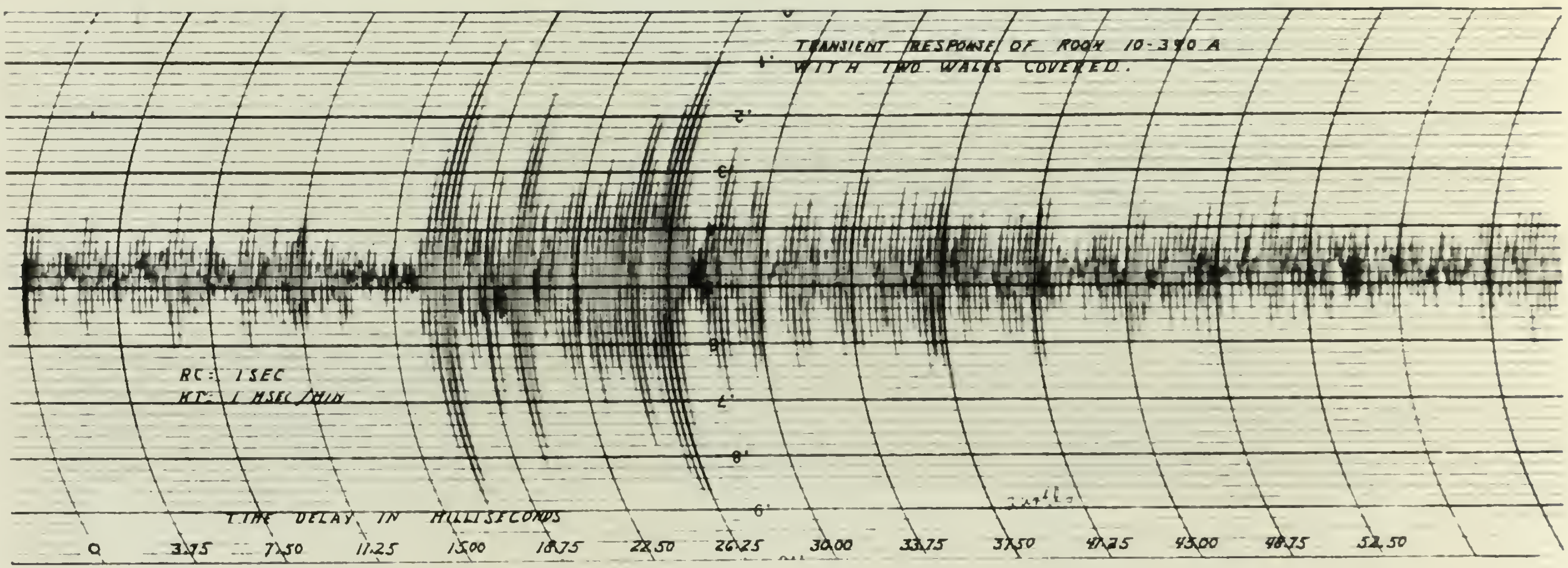

TRANSIENT RESPONSE OF ROOM 10-390 A
WITH TWO WALLS COVERED.
RC: 1SEC
TIME DELAY IN MILLISECONDS
0
3.75
7.50
11.25
15.00
18.75
22.50
26.25
30.00
33.75
37.50
41.25
45.00
48.75
52.50

Figure 4.13

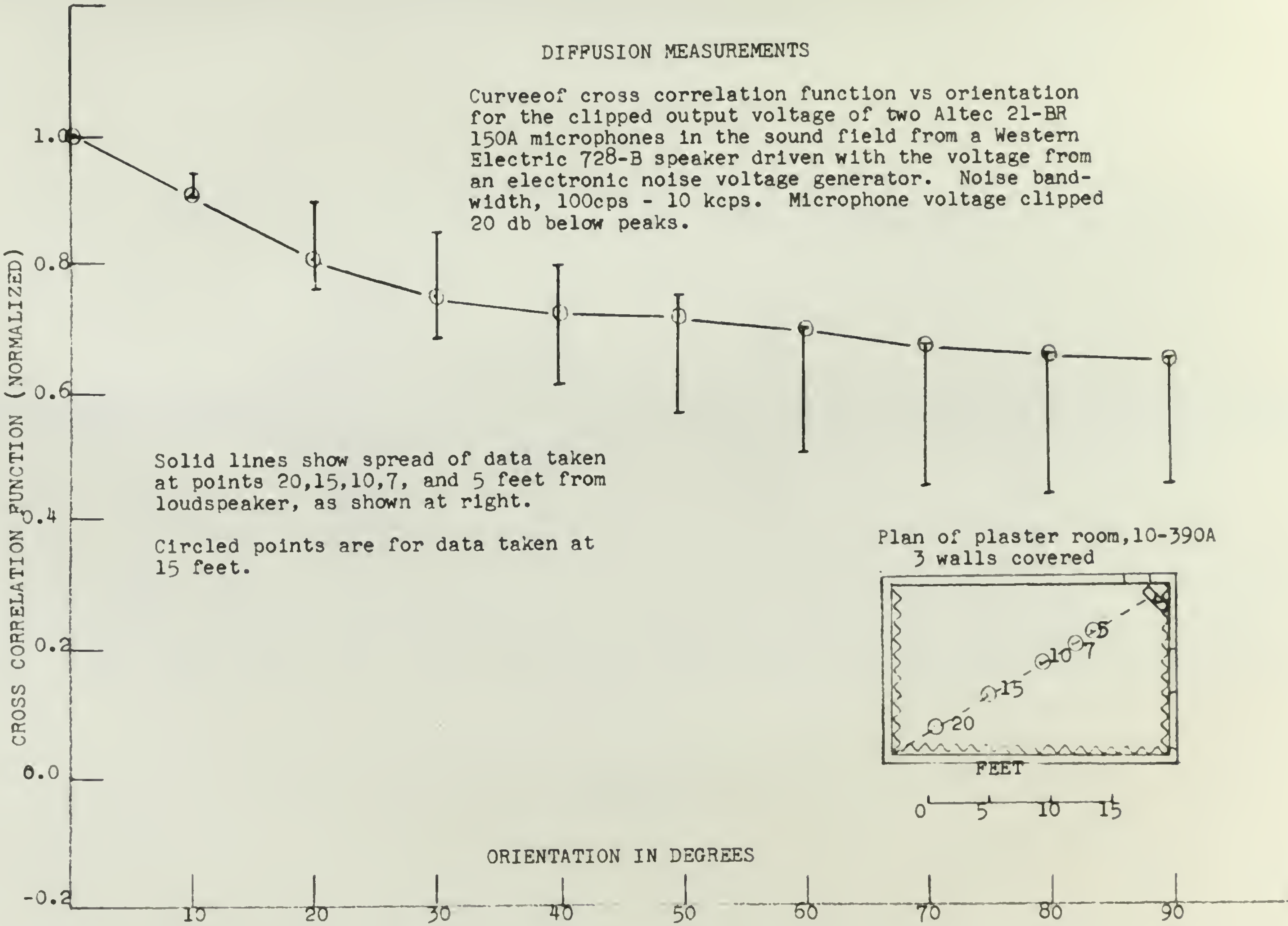

DIFFUSION MEASUREMENTS
Curveeof cross correlation function vs orientation for the clipped output voltage of two Altec 21-BR 150A microphones in the sound field from a Western Electric 728-B speaker driven with the voltage from an electronic noise voltage generator. Noise band-width, 100cps - 10 kcps. Microphone voltage clipped 20 db below peaks.
Solid lines show spread of data taken at points 20,15,10,7, and 5 feet from loudspeaker, as shown at right.
Circled points are for data taken at 15 feet.
Plan of plaster room,10-390A
3 walls covered
FEET
0
5
10
15
20
7
CROSS CORRELATION FUNCTION (NORMALIZED)
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
ORIENTATION IN DEGREES
10
20
30
40
50
60
70
80
90

Figure 4.14

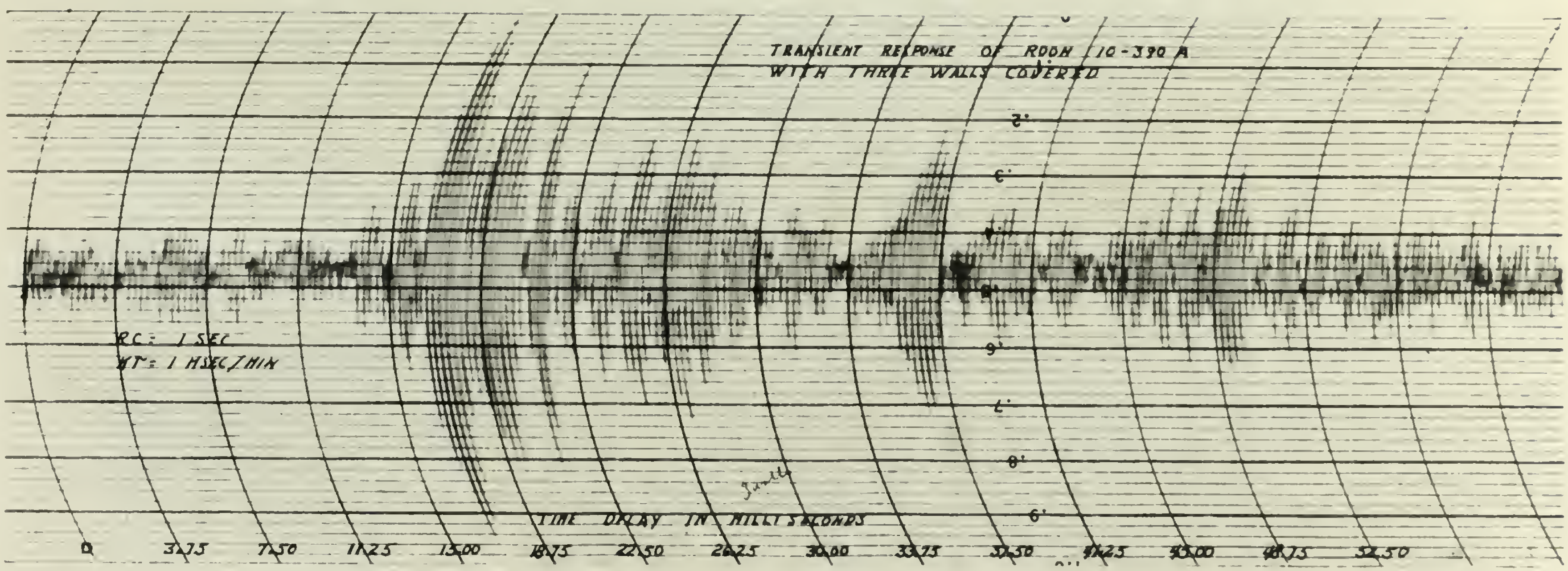

TRANSIENT RESPONSE OF ROOM 10-390 A
WITH THREE WALLS COVERED
RC: 1 SEC
WT: 1 MSEC/MIN
TIME DELAY IN MILLISECONDS
0
3.75
7.50
11.25
15.00
18.75
22.50
26.25
30.00
33.75
37.50
41.25
45.00
48.75
52.50

Figure 4.15

4.4 DISCUSSION OF EXPERIMENTAL RESULTS

In reviewing the experimental results of Section 4.3 it appears that the measuring system and the techniques employed yield results similar to those obtainable with a highly directional microphone. It appears that the practice of setting the microphones a certain distance apart corresponds to making a microphone which is extremely sensitive in measuring the direction of propogation for a sound having a particular specular composition. Generally, setting the microphones further apart makes the array more sensitive to spectra having lower geometric mean frequencies. Insufficient data is available to permit an evaluation of this method of measuring diffusion. However, the author believes that the data presented herewith is sufficient to indicate a possible experimental method of obtaining a measurement of the degree of diffusion of a sound field. Moreover, it is believed that this method warrants further study.

4.5 SUGGESTIONS FOR ADDITIONAL WORK

On the basis of the discussion and measurements of diffusion made in connection with this project, it is felt that there is some potential merit in the techniques employed. As a suggestion for additional work, it is proposed that some consideration be given to further diffusion

measurements made with various bandwidths of noise voltage. Such a project would make extensive experiments in some room such as the model studio. One possible series of experiments would involve measurements for various microphone separations (selected in a manner similar to that of Experiment 1, Section 4.3) and various conditions of diffusion (adjustments in diffusion to be made by rearranging variable diffusers in the model studio).

V. CONCLUSION

5.1 SUMMARY OF CORRELATION MEASUREMENTS

Correlation measurements as studied in this project have been of two types - (1) measurements of the transient response of rooms and (2) measurements based on the definition of a diffuse sound field. It is an interesting coincidence that both types of measurements can be made with practically the same experimental setup. Transient response data may be taken by cross correlating between the microphone output voltages of two microphones, one of which is located near the source and the other in the reverberant field of the room. Diffusion measurements are made in the same way, by cross correlating between the microphone output voltages of two microphones which are spaced a particular distance apart and oriented in a certain way with respect to the source.

The above illustrated coincidence serves to point out the versatility of correlation techniques. Unfortunately, this coincidence also serves to illustrate one of a number of pitfalls that one may encounter in the use of such techniques. Improper microphone placement in either type of measurement would lead to spurious results.

It was stated before that diffusion measurements by correlation techniques seemed to yield results similar to those obtainable by use of a highly directional microphone.

It should be of interest to point out that spacing microphones a distance apart which is calculated from auto-correlation data corresponds to calibrating the microphone array for maximum directionality for the particular frequency spectrum of noise voltage being used.

This directive microphone array constructed by correlation techniques appears to have two major lobes in its directivity pattern. These major lobes are 180 degrees apart. The calibration process consists of spacing the microphones so as to eliminate any minor lobes in the directivity characteristics of the array.

It should be pointed out that the directivity which we have been talking about refers only to the ability of the microphone array and the associated correlation equipment to discriminate between the presence of a plane wave front which reaches both microphones at the same instant and a plane wave front which reaches each microphone at a different instant.

In a previously mentioned report by the British Broadcasting Company, it was concluded on the basis of their experimental studies that the methods of short pulse analysis are not satisfactory for the investigation of diffusion in full scale rooms. The pulse methods referred to in this conclusion involved - (a) studies of the irregularity of the envelope of the decay for a short

pulse of sound and (b) variation in these short pulse decay irregularities with position in the room.

Another group of investigators[13] using a directional microphone technique attempted to obtain a steady state measure of the diffuseness of sound fields. The conclusion to their research cited the necessity for a microphone that could be made highly directional over the whole audio frequency range.

It appears that the correlation method of measuring transient response is suitable for use in full scale rooms. Unlike short pulse methods, correlation analysis of transient response may be made to yield short term decay rates for various bandwidths of frequencies. On the basis of this latter fact, it may be concluded that correlation analysis presents a connecting link between room response data gathered by steady state methods and that gathered by pulse methods.

In addition, it can be seen that correlation techniques, when used in conjunction with a microphone array calibrated for the purpose, can be made to yield data similar to that obtained by directional microphone methods. Further, the microphone array can be calibrated for any bandwidth by use of data taken from its autocorrelation curve.

From these facts it may be further concluded that the correlation method of analysis of bounded sound fields is potentially more informative than either steady state methods or pulse methods taken alone.

APPENDIX

Reprinted from THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, Vol. 22, No. 3, 328–340, May, 1950

Pulse Statistics Analysis of Room Acoustics*

R. H. BOLT, P. E. DOAK, AND P. J. WESTERVELT
Acoustics Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received February 2, 1950)

Many sounds of speech and music more nearly resemble pulsed wave trains than abruptly terminated continuous sounds as used in reverberation measurement. It is therefore not surprising to find that two rooms can differ markedly in acoustical quality even if they appear identical under reverberation analysis which ignores details of short transients.

This paper introduces a pulse statistics point of view which takes immediate account of the pulse-like nature of common sounds. Fundamentally, the method consists in examining the response of the room to a short pulse. The walls are replaced by an array of image sources (simple images if the walls are hard, or appropriately modified if there is absorption). These image arrays are then considered statistically.

From this approach one can derive such classical quantities as reverberation time and mean free path. One can also analyze the detailed nature of discrete reflections including interference effects, and thus obtain an average correlation between room geometry and the character of its pulse response.

Idealized experiments in a hard-walled rectangular room are employed to illustrate the essential features of this approach. A point source emits an exponential damped 3600-c.p.s. wave train of about 2 msec. duration. The received signals are recorded logarithmically on an oscillograph and the system is calibrated for quantitative results. Several dozen discrete reflections can be measured and correlated with calculation. The pulses merge into a more or less continuous background after a time that is calculated and confirmed experimentally. Detailed differences arise according to the positions of the source and microphone in the room.

INTRODUCTION

It is well known that reverberation time is not a completely adequate index of the acoustical quality of a room. Implicit in the concept of reverberation time is the assumption that the room reaches a steady-state diffuse condition before the source of sound is abruptly terminated. In practice, however, most sounds of speech and music can be generally classified as pulsed wave trains whose amplitudes and frequency components fluctuate sufficiently within time intervals shorter than the time constant of the room, so that the room seldom reaches steady state. Thus it would seem that the response of the room to transient sounds of this general type is an especially important physical problem of room acoustics. The results obtained recently by Mason and Moir[1] and others[2] who have used short tone bursts to investigate acoustics of auditoriums lend support to this point of view.

The task of describing mathematically the response of a room to an arbitrary transient, and of studying the roles of room geometry and distribution of absorbing materials in this response, is extremely complicated. The problem can be approached from a normal mode point of view,[3] or one can attempt to "follow" the sound waves around in the room as they are reflected back and forth from the walls. The latter approach has been recently investigated by Mintzer[4] using Laplace transform methods.

From the transient point of view, it is desirable to use the second approach. In essence, this method consists of replacing the effect of the boundaries of the room by an infinite array of image sources, each image corresponding to one of the multiple reflections of the original wave emitted by the source. Finding these images analytically is no simple matter, and only in very special cases will the images be "mirror images" of the source.[3] The image concept has been used extensively in earlier geometric studies[5,6] where the source is considered to be incoherent. Due mainly to mathematical difficulties, little use of images has been made as yet in wave-acoustical investigations of rooms.

However, it has become increasingly apparent in recent years that the first 20 db of decay of a sound in a room is of primary importance in differentiating between two rooms which have approximately equal overall reverberation times. Further, as the work of Mason and Moir indicates,[1] the time and amplitude distributions of reflected tone bursts can be correlated with the acoustical quality of a room. These facts indicate that the images relatively close to the source, i.e., the first few reflections, are primarily responsible for certain important features of the acoustical character of rooms, as Brillouin has observed.[6] It should therefore be worth while to study this "short term" transient response by a method of images in which all wave properties of the image sources can be considered (i.e., where the assumption of an incoherent source is not made). Further, if an image array satisfying the boundary conditions can be found, one should be able to treat this array statistically and thus obtain the long term average transient response as well.

This paper is confined for the most part to a discussion of an idealized case, a hard-walled rectangular

* This work was supported in part by the ONR, Department of Navy, under Contract NObs 25391, Task 7.

[1] C. A. Mason and J. Moir, "Acoustics of cinema auditoria," J. Elec. Eng. 88, Part III, No. 3 (September, 1941).

[2] British Broadcasting Corporation, Engineering Division, Research Department Reports B.027 and B.035.

[3] P. M. Morse and R. H. Bolt, "Sound waves in rooms," Rev. Mod. Phys. 16, 117 (1944).

[4] D. Mintzer, "Transient sounds in rooms," J. Acous. Soc. Am. 22, 341 (1950).

[5] C. F. Eyring, "Reverberation time in 'dead' rooms," J. Acous. Soc. Am. 1, 217–241 (1930).

[6] J. Brillouin, "Sur l'acoustique des salles," Rev. d'Acoustique 1 (September–November, 1932).

room containing a simple source which emits a short pulse. Statistical properties of the image array are investigated to illustrate the image method and to provide a basis for future work on more general cases.

The limited applicability of results based on a simple mirror image picture has already been pointed out.[3] However, as the discussions of both Morse and Bolt[3] and Mintzer[4] indicate, a specular approximation is allowable in cases where the walls are not too soft, and, in any case, one can assume that the image of a simple source can be described analytically by an expansion in spherical harmonics, which is essentially a multipole expansion of the image source taken about the mirror image point. Thus, if the walls are not fairly hard, neither the simple mirror image nor the specular approximation can be used, and other suitable representations of the images must be found.

PULSE STATISTICS THEORY

1. Image Space and the Time Distribution of Reflected Pulses

With these restrictions in mind, let us now set up a working picture for pulse analysis. We shall consider a simple rectangular room with perfectly reflecting walls. The room has dimensions L_x, L_y, L_z, one corner being at the origin of a Cartesian coordinate system.

A sharp pulse of sound is emitted from a source in this room. The source is assumed to be a simple one, with spherically symmetric radiation. This point source may be located by the vector:

$$\mathbf{r}_s = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}, \tag{1}$$

as shown in Fig. 1. A point receiver of sound pressure is located at a position:

$$\mathbf{r}_R = U\mathbf{i} + V\mathbf{j} + W\mathbf{k}. \tag{2}$$

The vector displacement of the receiver from the source is:

$$\mathbf{R}_0 = \mathbf{r}_s - \mathbf{r}_R. \tag{3}$$

Associated with this sound source is an infinite array of image sources each occupying, alone, an image replica of the physical room. This image array also is illustrated in Fig. 1. Each image cell is designated by three numbers (l, m, n), and these three numbers can take all integral values from minus infinity to plus infinity. The cell (0, 0, 0) is the actual room.

The pulse associated with the image (1, 0, 0) reflects once from the wall $x = L_x$. The pulse from (2, 0, 0) reflects first from the wall $x = 0$, then from the wall $x = L_x$. Obviously the absolute value of the cell number, l, for cells lying along the x axis gives directly the number of wall reflections suffered by the pulse from the image in question. The pulse from (1, 1, 0) reflects once from the wall $x = L_x$ and once from the wall $y = L_y$, so that its number of reflections is $|l| + |m|$. In fact, by this system of designation, the total number of wall reflections suffered by the pulse from the image (l, m, n) is directly:

$$N_{lmn} = |l| + |m| + |n|. \tag{4}$$

The vector position of each image source is:

$$\mathbf{r}_{lmn} = x_l\mathbf{i} + y_m\mathbf{j} + z_n\mathbf{k}. \tag{5}$$

The vector position of each image with respect to the receiver is:

$$\mathbf{R}_{lmn} = \mathbf{r}_{lmn} - \mathbf{r}_R. \tag{6}$$

To evaluate these last two equations, we note first that

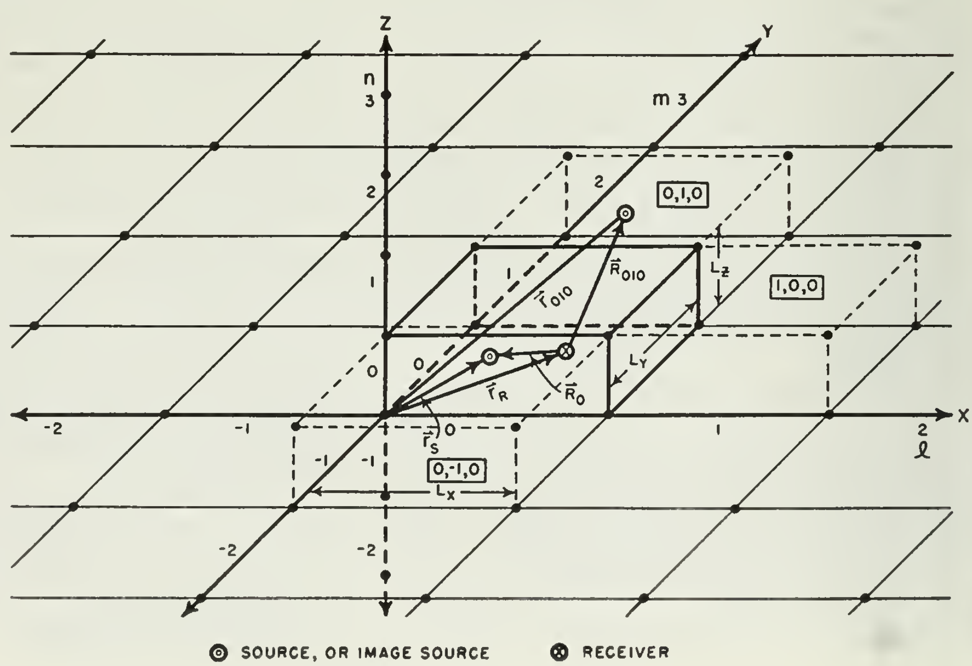


FIG. 1. An image array, showing image cells and source and receiver locations. The cell designation numbers are enclosed in the small boxes. ($L_x = 23.0$ ft., $L_y = 13.4$ ft., $L_z = 8.44$ ft.)

each image cell is a mirror reflection of the cells adjacent to it. This complicates the analysis somewhat in that we have two kinds of symmetry with respect to the basic coordinate system. Fortunately the cell designation numbers indicate directly the type of symmetry: even numbers, including 0 for the actual room, designate cells in which the source position duplicates that in the original room, while odd numbered cells have sources at a reflected position. Therefore the components for Eq. (6) are given by:

$$\left.\begin{aligned} x_l&=lL_x+X\\ y_m&=mL_y+Y\\ y_n&=nL_z+Z\end{aligned}\right\}\quad l,\,m,\,n\text{ even,}$$

$$\left.\begin{aligned} x_l&=(1+l)L_x-X\\ y_m&=(1+m)L_y-Y\\ z_n&=(1+n)L_z-Z\end{aligned}\right\}\quad l,\,m,\,n\text{ odd.}\tag{7}$$

In order to generalize some illustrative calculations we select the longest dimension of the room as a scale unit and introduce the following definitions:

$$\begin{gathered} L_x=L,\quad L_y=pL,\quad L_z=qL,\quad p,q\leq 1;\\ (U-X)/L=\mu_{xe},\quad (U+X)/L=\mu_{xo},\\ (V-Y)/L=\mu_{ye},\quad (V+Y)/L=\mu_{yo},\\ (W-Z)/L=\mu_{ze},\quad (W+Z)/L=\mu_{zo}.\end{gathered}\tag{8}$$

Thus p and q are dimension ratios and the μ's give the source-to-receiver displacements for both kinds of symmetry. From Eq. (6) we then get:

$$\frac{\mathbf{R}_{lmn}}{L}=\left\{\begin{matrix}[l-\mu_{xe}]\mathbf{i}\\ \text{or}\\ [l+1-\mu_{xo}]\mathbf{i}\end{matrix}\right\}+\left\{\begin{matrix}[mp-\mu_{ye}]\mathbf{j}\\ \text{or}\\ [(m+1)p-\mu_{yo}]\mathbf{j}\end{matrix}\right\}+\left\{\begin{matrix}[nq-\mu_{ze}]\mathbf{k}\\ \text{or}\\ [(n+1)q-\mu_{zo}]\mathbf{k}\end{matrix}\right\},\quad\left\{\begin{matrix}\text{even}\\ \text{or}\\ \text{odd}\end{matrix}\right\}.\tag{9}$$

Some calculations from this equation are illustrated in Figs. 2 and 3 which will be discussed later.

Next let us find the average number of pulses received up to a specified time, t, after the emission of the pulse from the original source. These pulses come from all of the images within a radius $|\mathbf{R}_{lmn}|=ct$. One image is contained in each cell of volume $V=L_xL_yL_z$. The number of pulses, N_p, is directly given by the volume of the sphere out to ct divided by the volume of one cell:

$$N_p=\frac{4\pi|\mathbf{R}_{lmn}|^3}{3V}=\frac{4\pi c^3t^3}{3V},\tag{10}$$

and the number of pulses received per second is:

$$\frac{dN_p}{dt}=\frac{4\pi c^3t^2}{V}.\tag{11}$$

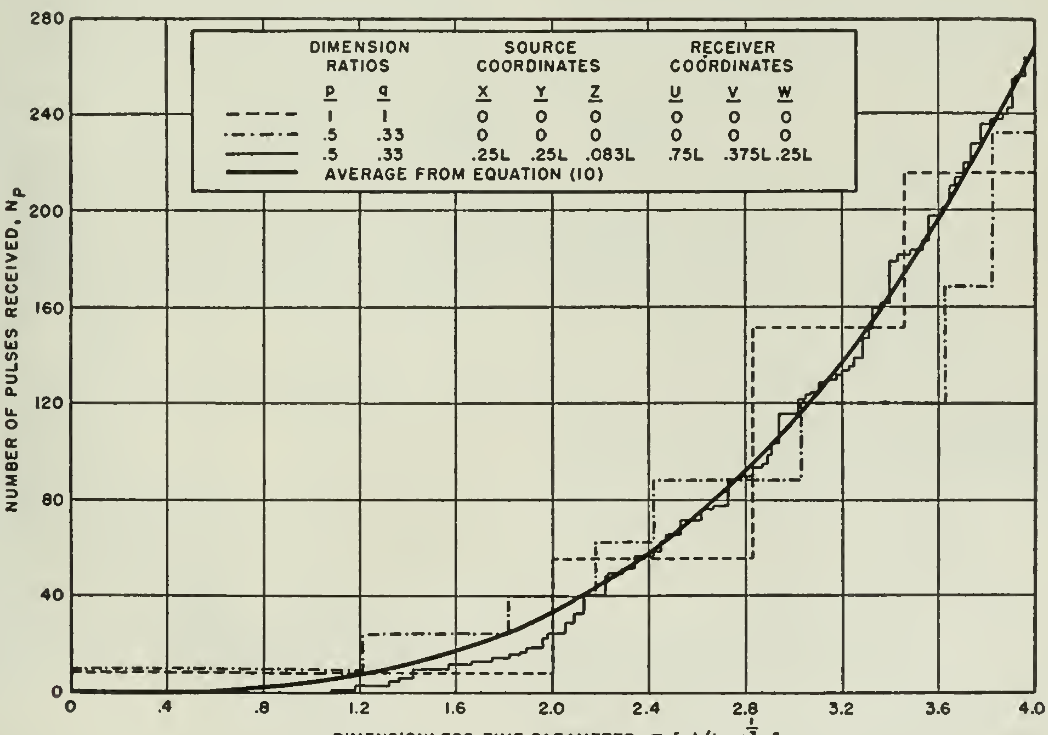


FIG. 2. Graph showing N_p, the number of pulses arriving at the receiver up to a time t as a function of the dimensionless time parameter τ. Note the extreme stepped behavior of the dashed curve for a cubical room.

These equations, (10) and (11), are valid for values of l, m, and n, large enough so that the details regarding positions of source and receiver in the room can be neglected. These equations also are illustrated in Figs. 2 and 3. A calculation of an actual case illustrates that a surprisingly large number of reflections per second are predicted. In a room of 10,000 cu. ft., Eq. (11) gives a rate of 180 pulses per second at 1/100 of a second. There are several conditions met in practice which greatly reduce this number. For one thing, the floor is generally quite absorptive when an audience is present, so that half of the spherical volume just assumed is essentially eliminated. Also there are usually a large number of degeneracies or coincidences which reduce the effective number of pulses, as we shall see.

Figures 2 and 3 show N_p, the number of pulses received up to a time t, for rooms of several dimension ratios and various positions of source and receiver. N_p is plotted for convenience as a function of the dimenless time parameter

$$\tau = \frac{ct}{(pq)^{\frac{1}{3}}L}, \tag{12}$$

where p, q, and L are defined in Eq. (8). The heavy solid line in both figures gives the average value of N_p as computed from Eq. (10).

In Fig. 2 the dashed curve shows the extreme degeneracy of a cubical room when both source and receiver are in a corner. The pulses arrive in large groups because of the symmetry of the image point lattice for a cubical room and the location of source and receiver. The dash-dot curve for a rectangular room with $p=0.5$ and $q=0.33$ shows similar but smaller groups of coincident pulses. Both these curves start at $N_p=8$ since the direct pulse and the pulses from the seven source-corner images reach the receiver simultaneously at $t=\tau=0$. The light solid curve is for the same rectangular room, source and receiver now being at point such that coincidences are more or less "accidental." We note that this curve follows the average curve very closely for $\tau>2.2$. For $\tau<2.2$, the curve lies consistently below the average curve. This is partly because the separation of source and receiver are such that no pulse arrives at the receiver until $\tau=1.1$. In detail, this initial part of the curve depends in a somewhat more involved way on the relative positions of the real and image sources and the receiver.

Figure 3 shows in more detail the initial rise of the curve just discussed. The dash-dot and dashed curves show similar initial rises for a room of $p=0.8$ and $q=0.6$. The same general behavior is shown by these curves, although both reach the average curve sooner. The duration of this initial departure appears to decrease as the room becomes more nearly cubical.

2. Some Special Cases of Coincidences

In Eqs. (10) and (11) it was assumed that all images are distinct, that is, that no two pulses arrive at the receiver at exactly the same time. However, if source and receiver are both placed in certain locations, it is possible for the pulses to arrive in groups, as is evident from Figs. 2 and 3. It is of interest to find the

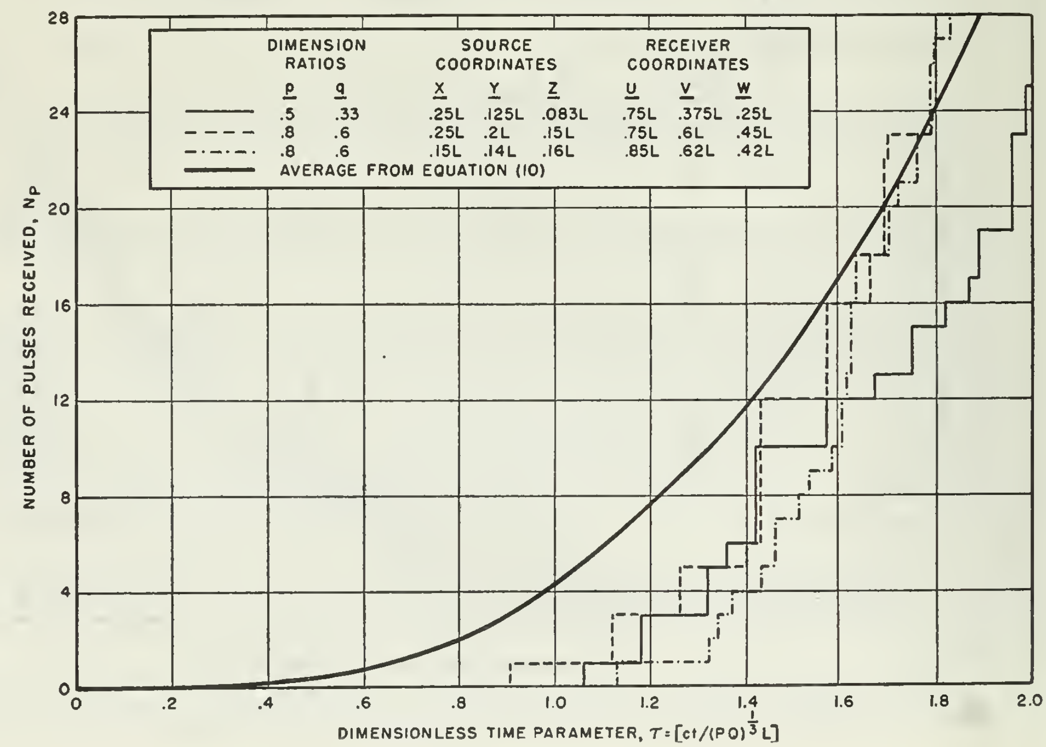


FIG. 3. Graph on an expanded scale showing the initial behavior of N_p as a function of τ.

number of distinct pulse groups reaching the receiver for some of these degenerate cases. In all cases it is assumed that the room is rectangular, hard-walled, and of incommensurate dimensions.

Case 1

Consider the source to be in a corner and the receiver in a position incommensurate with respect to the images but otherwise arbitrary. If the source is close enough to the corner, the images will clump together in groups of eight, being so close together that, effectively, each clump of eight acts as a single distinct image. Since pulses reaching the receiver will arrive in groups of eight, the *effective* number N_p' of distinct pulses arriving per second is:

$$N_p' = \frac{1}{8}\frac{4\pi(ct)^3}{3V} = \frac{\pi(ct)^3}{6V}. \qquad (13)$$

Case 2

Here we place both source and receiver in the center of the room. All the images are distinct, but because of the symmetry arising from the position of the source and receiver, a pulse from an image along a negative coordinate axis (the origin being taken at the center of the room for convenience) will arrive simultaneously with a pulse from an image along the positive coordinate axis. A pulse from a non-axial image in the first quadrant of a coordinate plane will arrive simultaneously with pulses from corresponding images in each of the other quadrants, and a pulse from an "oblique" image will arrive simultaneously with pulses from corresponding images in the other seven octants. Thus the effective distinct images are all contained in one octant of image space. Counting of the number of distinct pulses arriving can then be accomplished in exactly the same manner as the counting of the number of normal modes with frequencies less than a certain value.[3] The volume occupied in image space by all distinct image points within a radius, ct, divided by the volume, $V = L_xL_yL_z$, occupied by each image point, gives for N_p' in this case:

$$N_p' = \frac{\pi}{6V}(ct)^3 + \frac{\pi L_T}{32V}(ct)^2 + \frac{S}{8V}(ct), \qquad (14)$$

where:

$$V = L_xL_yL_z$$
$$S = 2(L_xL_y + L_yL_z + L_zL_x)$$
$$L_T = 4(L_x + L_y + L_z).$$

Similar arguments for the cases when both source and receiver are in a corner, at the center of a wall, or at the center of an edge show that all these cases (except *Case 1*, of course) can be expressed by the formula:

$$N_p' = \frac{1}{\eta_x\eta_y\eta_z V}\left[\frac{\pi(ct)^3}{6} + \frac{\pi}{8}(\eta_xL_x + \eta_yL_y + \eta_zL_z)(ct)^2 + \tfrac{1}{4}(\eta_x\eta_yL_xL_y + \eta_y\eta_zL_yL_z + \eta_z\eta_xL_zL_x)(ct)\right], \qquad (15)$$

where η_x, η_y, and η_z are to be chosen as follows:

$\eta_x = \eta_y = \eta_z = 2$,	source and receiver in corner;
$\eta_x = \eta_y = 2$, $\eta_z = 1$	source and receiver at center of a z-edge, etc.;
$\eta_x = 2$, $\eta_y = \eta_z = 1$	source and receiver at center of a yz-wall, etc.;
$\eta_x = \eta_y = \eta_z = 1$,	source and receiver at center of room.

It is clear that for a given room the formula for N_p' will depend upon the position of both the source and receiver. Further, for pulses of a finite length, for actual sources and receivers, and for most positions of source and receiver, exact coincidences will be rare, and "almost coincidences," with interference effects between the various pulses, will be the rule rather than the exception.

3. Pulse Spacing Statistics

The close similarity between the normal frequency lattice and the three-dimensional image source lattice has already been noted. The average statistical properties of pulses bear a close resemblance to the equivalent properties of the normal frequencies. Therefore, similarities in the fluctuation statistical properties common to both normal frequencies and pulses might be expected.

The frequency spacing index ψ has been defined as the mean squared ratio of actual to average normal frequency spaces, for a specified frequency interval. This index has been evaluated for a certain class of rectangular rooms, and it can be calculated if normal frequency values are available.[7] We shall see next that the exact analog of ψ for pulse spacing (defined as the mean squared ratio of actual to average time intervals for returning pulses) can be calculated and evaluated in every case for which the frequency spacing index index can be calculated or evaluated, provided that source and receiver are maintained in one corner.

We start with the expression for the dimensionless normal frequency of the l, m, nth mode:[7]

$$\mu_{lmn} = \tfrac{1}{2}(pq)^{\frac{1}{3}}[l^2 + (m/p)^2 + (n/q)^2]^{\frac{1}{2}}. \qquad (16)$$
$$l, m, n = 0, 1, 2\cdots.$$

We obtain the time of arrival of distinct pulses from the image sources by using Eq. (15) for the case where source and receiver are both in a corner. The resulting equation for time of arrival can be put into the dimensionless form,

$$\tau/4 = \tfrac{1}{2}(pq)^{-\frac{1}{3}}[l^2 + (mp)^2 + (nq)^2]^{\frac{1}{2}}, \qquad (17)$$
$$l, m, n = 0, 1, 2\cdots,$$

where τ is as defined in Eq. (12).

Notice that Eq. (17), which applies only to the case where source and receiver lie in the same corner, differs from Eq. (16) for μ_{lmn} only through the appearance of $(p)^{-1}$ and $(q)^{-1}$ in place of p and q. Since p and q enter

[7] R. H. Bolt, "Normal frequency spacing statistics," J. Acous. Soc. Am. 19, 79 (1947).

symmetrically in both Eqs. (16) and (17), they can be interchanged. Thus in order to convert any equation giving the frequency spacing index into a form which will yield the pulse spacing index, one need only substitute p^{-1} for q, q^{-1} for p, and $\tau/4$ for μ. Of course, if the room proportions are such that the required weighting factors cannot be evaluated, but actual pulse spacing values are available, for example from experiment, ψ for pulse spacing is easily evaluated from its defining equation:

$$\psi_{AB}=\frac{1}{\tau_B-\tau_A}\sum_A^B\left(\frac{\delta\tau_{lmn}}{\langle\delta\tau\rangle_{\text{Av}}}\right)^2\langle\delta\tau\rangle_{\text{Av}},\tag{18}$$

where $\langle\delta\tau\rangle_{\text{Av}}/4$ is obtained by making the substitutions $p\rightarrow(q)^{-1}$, $q\rightarrow(p)^{-1}$, $\mu\rightarrow\tau/4$ in the equation giving the average spacing between adjacent normal frequencies.[7]

4. A Derivation of the Mean Free Path

Pulse analysis yields a straightforward derivation of the classical mean free path. We assume that the order numbers are so large that the coordinates of the source and receiver can be neglected (that is, source and receiver can be considered to be at the origin). Using polar coordinates (r, θ, ϕ) we can express the number of reflections associated with each cell as:

$$N_{lmn}=r\left[\frac{\cos\phi\,\cos\theta}{L_x}+\frac{\sin\phi}{L_y}+\frac{\cos\phi\,\sin\theta}{L_z}\right].\tag{19}$$

We next obtain an average value of the number of reflections out to a given radius $r=ct$, by averaging Eq. (19) over all angles:

$$\bar{N}_{lmn}=\frac{\int_\sigma N_{lmn}d\sigma}{S_0}=\frac{rS}{4V},\tag{20}$$

where $d\sigma=r^2\cos\phi d\phi d\theta$, $S_0=(4\pi r^2)/8$, $S=2(L_yL_z+L_zL_x+L_xL_y)$, and the integration is taken over one octant of space. By definition, the mean free path is equal to the total distance traveled by an average pulse in a given time t, divided by the number of reflections of that pulse during the same time. Therefore:

$$\text{m.f.p.}=r/\bar{N}_{lmn}=4V/S.\tag{21}$$

5. Energy in the Pulse

We now consider the energy contained in a sound pulse and follow the course of that energy as the sound becomes dispersed throughout the room and absorbed at its boundaries. The total power radiated by a simple source is:[8]

$$\Pi=\frac{\pi\rho\nu^2Q_0^2}{2c}=\frac{4\pi}{\rho c}(\langle p_0^2\rangle r_0^2)\text{ ergs/sec.,}\tag{22}$$

[8] P. M. Morse, *Vibration and Sound* (McGraw-Hill Book Company, Inc., New York, 1948), second edition.

where Q_0 is the source strength, and $\langle p_0^2\rangle$ is the mean square sound pressure at a distance r_0 from the source in a free field. This equation applies to steady-state radiation. It is also valid for a single pulse wave train which has a sufficiently narrow spectral distribution. This requirement is fairly well satisfied, for example, if there are ten or more waves in the train, and if the wave amplitude is fairly constant throughout the duration of the pulse. We designate the length of this pulse τ_p and write the total energy contained in the pulse as:

$$E_p=\Pi\tau_p=(4\pi/\rho c)(\langle p_0^2\rangle r_0^2)\tau_p\text{ ergs.}\tag{23}$$

The energy density in the pulse is continually diminishing as the pulse radiates outward. At any instant of time, t, after emission, the volume of space containing the pulse is:

$$V_p=(4\pi/3)c^3[t^3-(t-\tau_p)^3]\simeq 4\pi c^3t^2\tau_p.\tag{24}$$

We next consider the multiplicity of pulses arriving at the receiver from all directions as time progresses (still assuming that the room is lossless). We take the number of pulses arriving per second, as given by Eq. (11), and multiply this by the energy density during the duration of a pulse passage, obtaining the energy density per second (provided the pulses add incoherently):

$$W_p=\frac{E_p}{V_p}\cdot\frac{dN_p}{dt}=\frac{4\pi}{\rho c}\frac{\langle p_0^2\rangle r_0^2}{V}\text{ ergs/cc-sec.}\tag{25}$$

If we now multiply W_p by the fraction of a second occupied by the individual pulse, we obtain the average energy density in a room of volume V:

$$W_R=\frac{4\pi}{\rho c}\frac{\langle p_0^2\rangle r_0^2}{V}\tau_p\text{ ergs/cc.}\tag{26}$$

A more convenient quantity experimentally is the mean square pressure in the room:

$$\langle p_R^2\rangle=\rho c^2W_R=\frac{4\pi c(\langle p_0^2\rangle r_0^2)\tau_p}{V}.\tag{27}$$

Suppose we take a pulse of $\tau_p=20$ msec. duration which generates a mean square pressure of one dyne (74 db pressure level) at a distance of one meter from the source in a free field. If this pulse is emitted in a room of 10,000 cu. ft. volume, we find that the average mean square pressure throughout the room after the pulse is dispersed, as calculated from Eq. (27), is about 69 db, or only 5 db less than the sound pressure level in the original pulse as measured at one meter. In practice, the sound is rapidly dissipated by absorption and may be canceled out by interference effects.

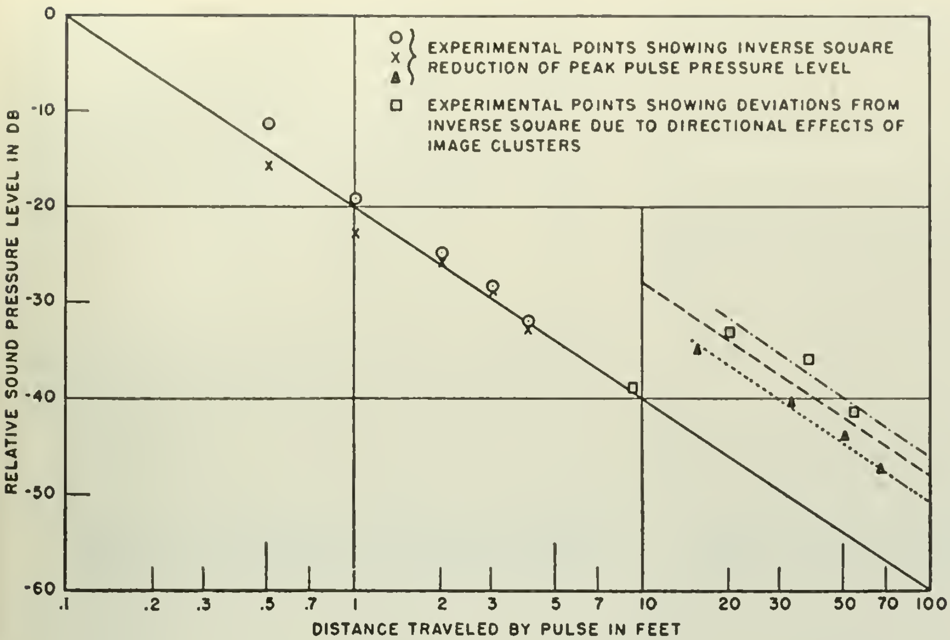


FIG. 4. Graphs of peak pulse pressure level as functions of distance from the source for several sequences of direct and reflected pulses.

6. A Derivation of the Eyring Reverberation Equation

Let us assume that all of the boundaries of the room are equally absorptive, and that their absorption is less than 20 percent, so that the specular reflection approximation is valid. Further, let us suppose that the energy in a pulse is diminished upon reflection by a factor $(1-\alpha)$, where α is the average over all angles of the free wave absorption coefficient. Then the pulse energy associated with the Nth image is:

$$E_{pN}=E_p(1-\alpha)^{\bar{N}}=E_p(1-\alpha)^{(ctS/4V)}, \qquad (28)$$

which leads at once to the Eyring reverberation equation:

$$T_{60}=KV/S\ \ln[1/(1-\alpha)].$$

This is the average long term reverberation time of the room. To study the short term response for a fairly hard-walled room, one could utilize the specular reflection approximation[3,4] for each of the first few reflections and thus obtain in detail the approximate short term response of the room to a particular transient. Further refinements of the long term response might be obtained by treating groups of images statistically, just as modified decay equations have been obtained by grouping of normal modes.[3]

PULSE MEASUREMENTS IN A HARD-WALLED RECTANGULAR ROOM

1. Experimental Procedure

In order to obtain some simple results which could be interpreted from a pulse statistics point of view, short sound pulses were produced experimentally in a hard-plaster-walled rectangular room, and photographs of the sound arriving at a microphone were made using a cathode-ray oscillograph. The source used was a W.E. 713A receiver unit feeding into a $\frac{3}{4}$-in. diameter brass tube 8 in. long packed with steel wool to present a high acoustic impedance to the diaphragm. The effective source was the end of the tube, which was small and could be considered approximately a simple source. The microphone was a W.E. 633A dynamic type, the output of which was passed through an ERPI RA-363 octave filter and through the amplifier section of a ERPI RA 277-F sound analyzer.

The output of the analyzer was attenuated logarithmically by a Kay Labs Type 510-A Logaten and then put across the vertical deflection plates of a DuMont 247 cathode-ray oscillograph. A pulsed carrier of about 2 msec. duration was produced by mechanical switching of the speaker input which consisted of a 3600-c.p.s. signal from a Hewlett Packard 200 D audio oscillator amplified through a Fairchild audio amplifier. A slow (approximately 42 cm/sec. on the screen) external single sweep on the oscillograph was activated mechanically shortly before the beginning of the pulse. Thus the logarithm of the acoustic signal, picked up by the microphone during the 0.25-sec. interval after the pulse was emitted, was recorded linearly on the oscillograph screen as a linear function of time, and was photographed.

2. Calibration and Auxiliary Data

The horizontal sweep speed was calibrated by a 30-c.p.s. signal direct to the vertical plates from the oscillator. The vertical deflection was calibrated in 5-db steps, using the attenuator pad on the ERPI analyzer. The horizontal sweep was found to be linear within experimental error, while the vertical deflection was

found to be linear in db with a maximum deviation of ± 1 db over the main working range of 50 db. With somewhat larger variations from linearity at low amplitudes, the vertical deflection is approximately linear over a range of 70 db, which includes background level in the pictures.

The following assumptions and conventions were adopted in interpretation of the data: (a) All pictures were calibrated in db *vs.* ft., since time in the room corresponds to distance in image space; (b) it was assumed that the acoustic pulse emitted by the source in all pictures was the same in both shape and height; (c) a decibel reference level was established in terms of the peak amplitude and was taken as zero db at 0.1 ft. from the source, all other levels thus coming out negative; (d) the velocity of sound was taken as 1120 ft./sec.

In order to get a rough check on the validity of the assumption that the actual source was a simple source, two sets of pictures were taken in which source and receiver were placed near the center of the room (so that direct pulses would arrive without interference from images) and were spaced several different distances apart from 0.50 to 4.0 ft. The peak levels of these direct pulses plotted against distance between source and receiver on semi-log paper are shown in Fig. 4. The best straight line fit to the experimental points is a line with inverse square slope. This in turn indicates that interpreting the actual source as a simple point source is a fair approximation. Another check on this approximation was made by measuring peak pulse amplitude in db at a constant radius from the source, but at six different angles with respect to the tube, from 0° to 180°. These measurements showed a maximum variation of ± 2 db in the peak values.

3. Pulse Shape and Energy

Pulse length and carrier frequency were selected empirically. The pulse length was chosen short enough compared to room dimensions to give at least several clearly separated "echoes," but long enough to include at least five or ten cycles of the carrier. In turn the carrier frequency was selected as high as possible

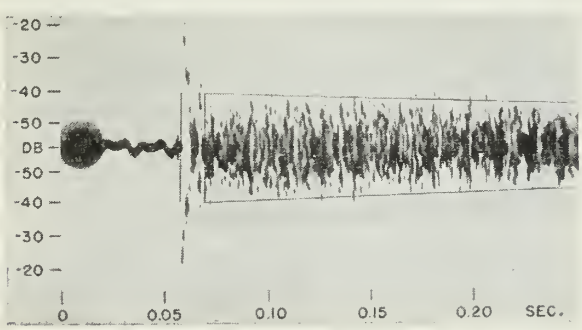


Fig. 5. Photograph for a non-degenerate case in which the distance between source and receiver is 1′. This photograph shows a typical direct pulse without interference from image pulses.

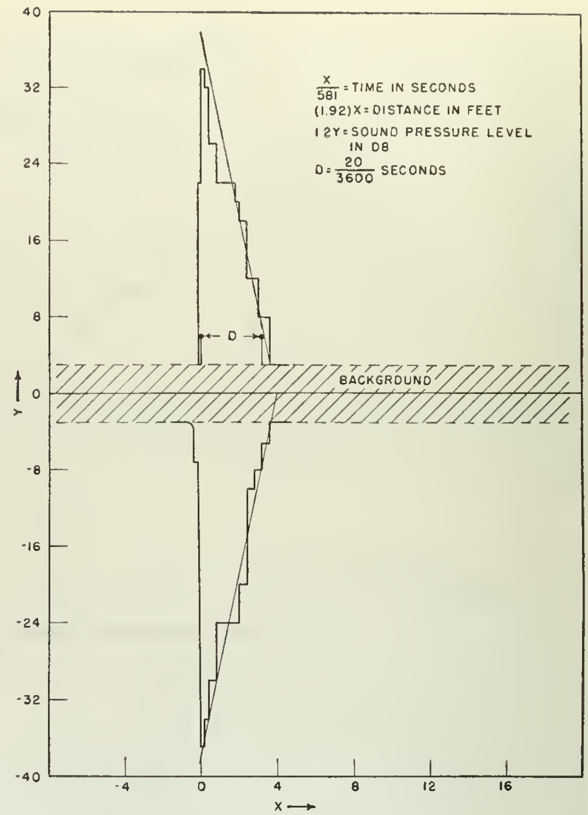


Fig. 6. Graphical enlargement of the direct pulse shown in Fig. 5, illustrating the general shape of the envelope.

(without encountering large air absorption) in order that the speaker would operate in a region of high efficiency and good transient response. The sweep speed and carrier frequency were such that carrier details are barely unresolvable, although individual oscillations of the carrier can be distinguished in some places in the pictures.

The exact nature and shape of the pulse was subsequently determined by analysis of the pictures (e.g., Fig. 5). The average pulse dimensions as determined from the photographs are illustrated in Fig. 6, which is a graphical enlargement of the direct pulse shown in Fig. 5. The detailed shape of the pulse is represented by a stepped curve in which the width of most of the steps is equivalent to several cycles of the carrier. At the onset of the pulse the first two or three swings of the alternating and exponentially increasing carrier are just discernible and the pulse builds up to maximum amplitude in about two cycles of the 3600-c.p.s. carrier. The pulse is almost symmetric but its peak amplitude in the positive direction is measurably greater than in the negative direction for this particular picture. (Symmetry of the pulse, of course, depends upon the phase of the carrier at the instants of switching it on and off.) Following the peak, the pulse decays in a rigorously exponential manner within experimental error as indicated in Fig. 6 by the two straight lines drawn through

the stepped curves. The pulse decay rate is about 2.5×10^3 db/sec., and depends principally upon the loudspeaker characteristics. The background noise showing on Fig. 5 is mainly a.c. power ripple. Maximum background level for all pictures was -65 db.

The energy in a pulsed carrier having an envelope that rises instantaneously and then decays exponentially is

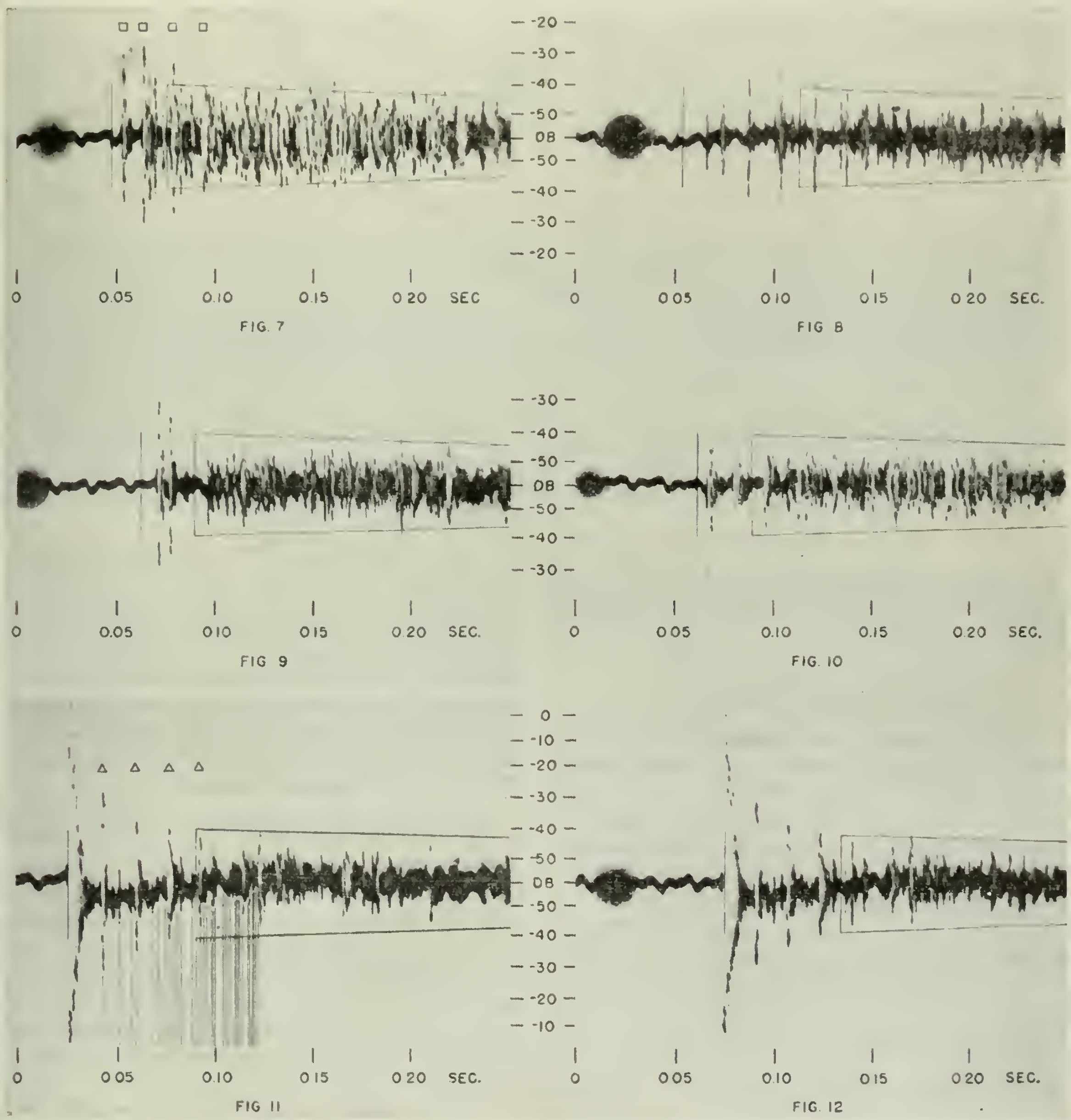
$$E_p=\frac{\pi p_0^2 r_0^2}{\rho c}\frac{\omega^2+2\gamma^2}{\gamma(\omega^2+\gamma^2)}\text{ ergs,}\tag{29}$$

where p_0 is the peak pressure in dynes at r_0 cm, ω is the angular frequency, and γ is the exponential decay constant. For the pulse shown here $\gamma=580$ and $\omega=7200\pi$ so that $(\gamma/\omega)^2\ll1$. Hence Eq. (29) reduces to

$$E_p=\frac{\pi p_0^2 r_0^2}{\rho c\gamma}.\tag{30}$$

If we let p_s be the r.m.s. pressure in the room a relatively long time after emission of the pulse (i.e., when the sound is fairly uniformly distributed), then the ratio p_s/p_0 is, using Eq. (27):

$$\frac{p_s}{p_0}=\frac{\langle\langle p_R^2\rangle\rangle^{\frac{1}{2}}}{p_0}=\frac{1}{p_0}\left(\frac{\rho c^2 E_p}{V}\right)^{\frac{1}{2}},$$



Figs. 7–12. Series of pulse photographs showing response in a hard-walled rectangular room when the source is placed in a corner and the receiver is located at various points around the room. Note the extreme number of coincidences when the receiver is also in a corner (e.g., Figs. 8 and 11).

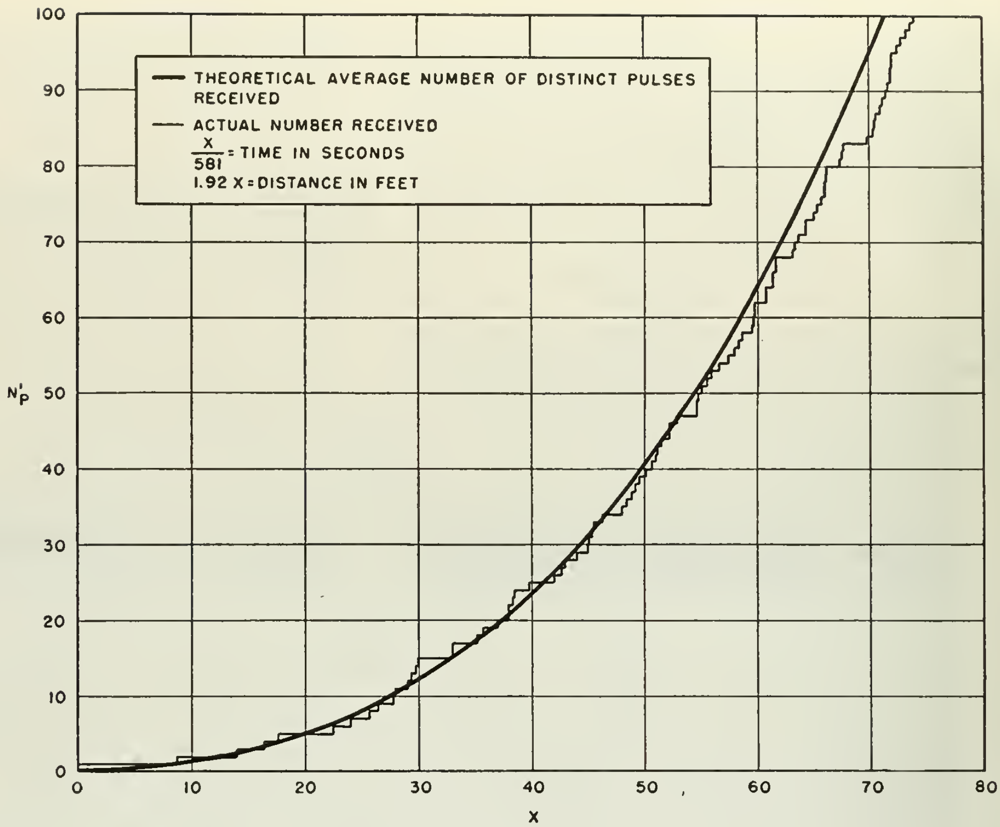


FIG. 13. The total number of distinct pulses N_p' arriving at the receiver, plotted as a function of time, when the source and receiver are both in the corner of the room, as in Fig. 11.

or, using Eq. (30) and letting $\tau_p - \frac{1}{2}\gamma$:

$$\frac{p_s}{p_0} = \left(\frac{2\pi r_0^2 c\tau_p}{V}\right)^{\frac{1}{2}}. \tag{31}$$

This expression neglects dissipation of energy in the room which is easily taken into account if the reverberation time of the room is known.

4. The Period of Resolution

A period of resolution t_r can be defined as the time at which the expected interval between successive pulses is just equal to the effective pulse duration τ_p. At a time t_r after the first pulse, one would no longer expect to "see" the individual echoes in the clear, but rather a smear characterized by an envelope above which the peaks of the pulses would occasionally appear.

An expression for the period of resolution t_r in terms of the pulse width τ_p can be obtained from the correct (i.e., appropriate for the existing number of degeneracies) expression for N_p or N_p' given in Eqs. (10), (13)–(15) by solving the difference equation for t_r:

$$N_p(t_r) - N_p(t_r - \tau_p) = 1. \tag{32}$$

This has been done for the conditions of the experiments performed and leads to the following values for the period of resolution:

Source and receiver in arbitrary positions:	$t_r = 9.98 \times 10^{-3}$ sec.
Source in corner, receiver in arbitrary position:	$t_r = 26.3 \times 10^{-3}$ sec.
Source and receiver in corners:	$t_r = 56.8 \times 10^{-3}$ sec.

These values correspond to an effective pulse duration of $\tau_p = 1.74 \times 10^{-3}$ sec., which is the pulse width at an amplitude equal to $1/e$ times the peak amplitude. The resolving time as defined is depicted in each of the pulse pictures by a vertical line appearing, in every case but one, to the right of the initial pulse. The individual pictures will be discussed in detail later.

5. Envelope and Decay of Unresolved Sound

The height of the vertical line that indicates the resolving time has been adjusted to equal the r.m.s. sound level relative to the peak pulse level as obtained from Eq. (31). The subsequent decay of this more or less continuous sound is portrayed by the two slightly sloping horizontal lines that form the envelope of the unresolved sound. The slopes of the envelope were determined from the measured reverberation time of the room at 3600 c.p.s.

It will be noted that the size of the envelope is, in a number of cases, considerably too large (e.g., Figs. 8, 10, 11). In every case where the theoretical envelope does not fit the photograph, the source, or both the

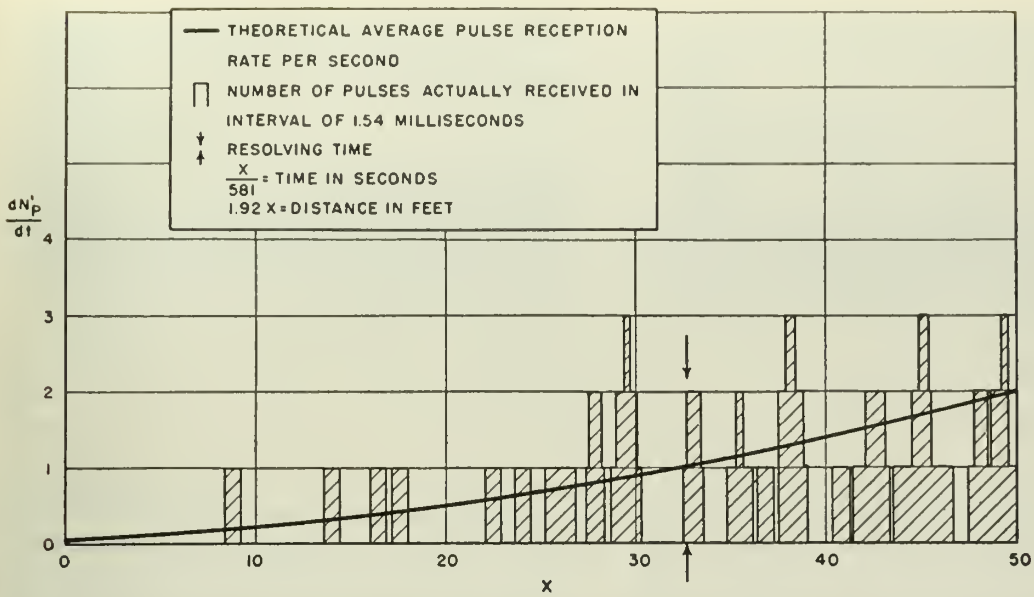
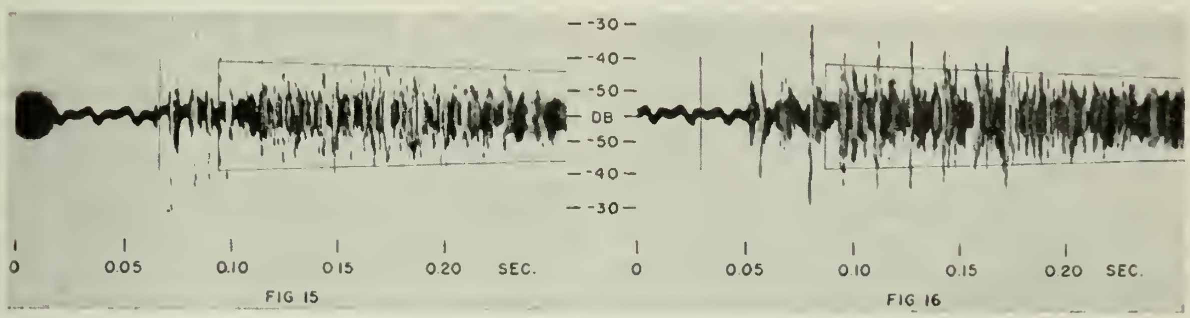


FIG. 14. The average and actual rates of arrival of pulses at the receiver when both source and receiver are in the corner of the room as in Fig. 11.



FIGS. 15 AND 16. Pulse responses for the receiver in other locations.

source and receiver, was in a corner position. Since the source was of finite dimensions it was never exactly in a corner but usually from 0.2 to 0.5 of a wave-length from one or more of the walls. Hence there was interference between the source and its images, resulting in a strongly directional *effective* source, and, consequently, in directional image clusters. Thus for the source in a corner and a particular location of the receiver, certain image sources will contribute very little effective energy to the receiver. This amounts to saying that the source effectively puts less energy into the room. This point will be made more evident from the discussion of the photographs.

6. Discussion of Pulse Photographs

The photographs were taken in a room of dimensions $L_x=23.0$, $L_y=13.4$ and $L_z=8.44$ ft. To specify the location of the source and receiver, we shall use (X, Y, Z) for the source position and (U, V, W) for the receiver position, distances being measured in feet.

In Fig. 5, the source coordinates are (11, 7.2, 3.9 ft.) and the receiver was located at (12, 7.2, 3.9 ft.). This position is non-degenerate and so the resolving time shown is $t_r=9.98$ msec. The envelope is seen to be slightly too large. This can perhaps be accounted for by the fact that both receiver and source are near the center of the room so that images from opposite pairs of walls interfere with each other consistently.

For Fig. 7, (XYZ) are (1, 1.5, 0.5 in.) and (UVW) are (5.75, 3.35, 2.11 ft.). The source is thus effectively eightfold degenerate, being in the corner, so the resolving time is $t_r=26.3$ msec. It will be noted that the direct pulse is exceeded in height by two succeeding pulses. The peak amplitudes of the pulses indicated by the squares are plotted on Fig. 4. This sequence corresponds to pulses from the successive images at (0, 0, 0), $(0, 0, 2L_z)$, $(0, 0, 4L_z)$, and $(0, 0, 6L_z)$, each of which is a cluster of eight image sources. The differences in amplitude show clearly that each image cluster has distinct directionality, with especially great variations in the X, Y plane. The long-dashed line on Fig. 4 is the peak level one would expect for incoherent addition of 16 sources (i.e., 12 db above peak for a single source). It is seen that the direct sound is 12 db below this line. However, the sound from the image sources lies along a line 14 db above the peak for a single source, which is,

in turn, 4 db below the level expected if the eight sources in each cluster all added coherently.

Figures 8–12 are a series in which (XYZ) are still (1, 1.5, 0.5 in.) and in which $U=1$ in., $W=1$ in. for all photographs. In Fig. 11, $V=L_y-1$; in Fig. 12, $V=3L_y/4$; in Fig. 13, $V=L_y/2$; and in Figs. 14 and 15, $V=1$ in. The directionality of the image clusters is clearly evident in Fig. 8. Here there is no distinct "direct" pulse since the pulse from the image at $(0, 2L_y, 0)$ arrives simultaneously with the pulse from the actual source. Evidently these two pulses interfere since the height of the first pulse is small. The series of tall pulses is again from the images at $(0, 0, 2L_z)$, $(0, 0, 4L_z)$ etc., each of which arrives simultaneously with corresponding pulses from the images at $(0, 2L_y, 2L_z)$, $(0, 2L_y, 4L_z)$, and their reflections in the X, Y plane. Each of these pulses then represents the simultaneous contributions of 32 image sources. As is evident by comparison of the peak heights with those in Fig. 7, the sources do not add coherently, however. The marked effect of interference is thus responsible for the low value of the actual envelope in this picture. The symmetry of the images corresponds closely to the symmetry when both source and microphone are in the same corner so the resolving time indicated is $t_r=56.8$ msec.

Figures 9 and 10 show effects similar to those just discussed for the other positions of the microphone mentioned previously. The resolving time for these is again $t_r=26.3$ msec. since the microphone position is no longer "degenerate." The influence of image-source directionality continues to be apparent.

In Figs. 11 and 12 both source and receiver are in the corner. Figure 12 serves to indicate the reproducibility of Fig. 11 which was possible with the experimental arrangements employed. The resolving time here is $t_r=56.8$ msec. The triangles refer to the points shown on Fig. 4. These pulses each represent the simultaneous contributions of the images at $(0, 0, \pm 2L_z)$, $(0, 0, \pm 4L_z)$ etc. As is evident from Fig. 4, the peaks very nearly follow the inverse square law. The peak amplitude, however, corresponds to 9 db above peak for a single source, whereas each pulse comes from 16 sources. The resulting peak in this case is less than would be expected for incoherent addition. Consequently it is again not surprising that the actual envelope is below the predicted value. The lines drawn on the lower half of the photograph indicate the calculated times of arrival of distinct pulses. Figure 13 shows the total number of distinct pulses N_p' arriving up to a time t, as calculated from Eq. (15) for this location of source and receiver; and Fig. 14 shows the average rate of arrival, and a graphical portrayal of the actual pulses arriving. Each line on Fig. 11 corresponds to a pulse on Fig. 14. It is seen that agreement is quite good.

In Fig. 15 the source is still in the corner but the receiver is at $(\frac{1}{4}L_x, \frac{1}{4}L_y, 1$ in.). This picture should be compared with Fig. 10.

In Fig. 16 the source remains at (1, 1.5, 0.5 in.) while the receiver is placed in the opposite corner (nearly) at $(L_x, L_y, 0)$. Here the pulse from the source arrives almost simultaneously with the pulses from the three image clusters at $(0, 2L_y, 0)$ $(2L_x, 2L_y, 0)$, and $(2L_x, 0, 0)$. The resulting interference is clearly evident from the appearance of the first pulse. In this case the symmetry is again almost the same as that for both source and receiver in the same corner. Therefore $t_r=56.8$ msec. is used. Agreement of actual and calculated envelopes is seen to be fairly good.

Although the present experiments are restricted to an idealized room, a close connection can already be seen between these results in a small room and pulse studies in large auditoriums. For example, in Fig. 17 the source is at (1.5, 6.7, 1 ft.) and the receiver at $(\frac{3}{4}L_x, \frac{1}{4}L_y, 9$ in.). These locations correspond to typical locations of a speaker and listener in a large rectangular hall. If we consider a scaling factor of about 5, the dimensions of this hall would be $115\times 67\times 42.2$ ft., and our pulse would have a corresponding width (at $1/e$ peak amplitude) of 8 msec. The scaled carrier frequency would be about 700 c.p.s. The resolving time is seen to be shorter than the time actually required for the direct pulse to reach the receiver, and the character of the response shows that a "smear" sets in immediately. The actual and calculated envelopes are seen to agree well in this case, since no degeneracies are present.

Figures 18 and 19 can be compared qualitatively with Fig. 17. These were obtained in a motion picture theater of about 250,000 cu. ft. volume. Pulses were produced by applying a short pulsed carrier signal to the theater loudspeaker. In Fig. 18 the microphone was at a seat in the side section of the main floor about two-thirds of the way back from the stage, under the balcony. The pulse length was 5 msec., and the carrier frequency 2800 c.p.s.

In Fig. 19 the microphone was in the center section of the main floor, well in front of the balcony and somewhat off the center line. The pulse length was again 5 msec. and the carrier frequency was 1500 c.p.s. Qualitative observations indicated that the "hour glass" bulge in Fig. 19 can be correlated subjectively with a "slap echo" from the hard but irregular rear wall. This bulge

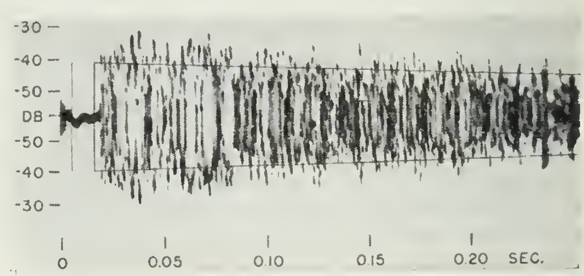
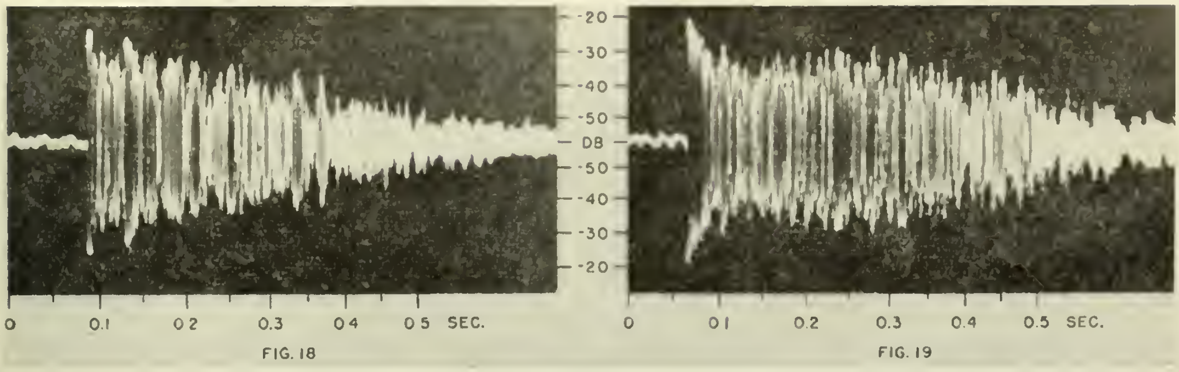


Fig. 17. Pulse response for non-degenerate locations of source and receiver which were chosen to correspond respectively to the positions of a speaker and listener in a large rectangular hall.



FIGS. 18 AND 19. Typical pulse responses in a motion picture theater of about 250,000 cu. ft. volume. Note qualitative similarity with Fig. 17.

was characteristic of pulse responses at this location for nearly all carrier frequencies. A more detailed analysis of these and similar photographs is in progress in an attempt to correlate various features of the pulse response with the results of subjective listening tests.

CONCLUSIONS

Most of the theoretical and experimental work presented in this paper is directly applicable only to hard-walled rectangular rooms. However, as has been pointed out, appropriate modifications of the images may be made when the walls are not hard and in some cases modifications for the first few images do not involve too great computational difficulties.

It appears that fairly detailed experimental and theoretical investigations of the short term transient response (e.g., the first 100–200 msec.) for rooms having various shapes and absorptive treatments should prove highly instructive. At the same time, the statistical methods outlined in this paper can be applied to the analysis of the fluctuation characteristics of the long term response.

The experimental results presented show that appropriate transients can be easily produced and observed, and, in simple cases at least, correlated with boundary conditions. Also it appears that pulse methods are readily adaptable to scale model experiments.

ACKNOWLEDGMENT

The authors wish to thank Dr. David Mintzer for permission to read his paper prior to publication and for participating in several discussions.

BIBLIOGRAPHY

1. Guilford, J.P. Psychometric Methods. New York: McGraw-Hill Book Company, Inc. pp. 6-7, 404-405.

2. Held, A. Statistical Theory With Engineering Applications. New York: John Wiley and Sons, Inc. pp. 662-665.

3. Wiener, N. The Extrapolation, Interpolation and Smoothing of Stationary Time Series. New York: John Wiley and Sons, Inc., 1949.

4. Faran, J.J. and Hills, R. "Correlators for Signal Reception" Acoustics Research Laboratory, Harvard University, Cambridge, Mass. September 1952.

5. Goff, Kenneth W. "The Application of Correlation Techniques to Some Acoustic Measurements," Doctor's Thesis, Electrical Engineering Department, M.I.T., Cambridge, Mass. (1954).

6. Sabine, W.C. Collected Papers on Acoustics. Harvard University Press, 1922.

7. Beranek, L.L. "Developments in Studio Design," Proceedings of the I.R.E., Vol. 38, No. 5, May 1950.

8. Bolt, R.H., Doak, P.E., and Westervelt, P.J., "Pulse Statistics Analysis of Room Acoustics," J. Acoust. Soc. Am., Vol . 22, No. 3, pp. 328-340, May 1950.

9. Mason, C.A., Moir, J., "Acoustics of Cinema Auditoria," J. Elec. Eng., Vol. 88, Part III, No. 3, September 1941.

10. Goff, Kenneth W. "Correlation Analysis," Research Notebook No. 1 (unpublished), M.I.T. Acoustics Laboratory.

11. Morse, P.M., and Bolt, R.H., "Sound Waves in Rooms," Review of Modern Physics, Vol. 16, No. 2, April 1944.

12. Bolt, R.H., and Roop, R.W., "Frequency Response Fluctuations in Rooms," J. Acoust. Soc. Am., Vol. 22, No. 22, pp. 280-289, March 1950.

13. Narain, C.K., Fay, R.D., and Bolt, R.H., "Study of the Distribution of Sound in Rooms," M.I.T. Acoustics Laboratory Quarterly Progress Report, April - June 1953.

Thesis
G457 Glass 28831
A correlation analysis of
bounded sound fields.